Journal of Geophysical Research: Oceans

RESEARCH ARTICLE
10.1029/2017JC012990

Special Section:
The Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) Project: Technologies, Methods, and Early Results

Observing the Ice-Covered Weddell Gyre With Profiling Floats: Position Uncertainties and Correlation Statistics

Paul M. Chamberlain[1], Lynne D. Talley[1], Matthew R. Mazloff[1], Stephen C. Riser[2], Kevin Speer[3], Alison R. Gray[4], and Armin Schwartzman[5]

[1]Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA, [2]College of the Environment, University of Washington, Seattle, WA, USA, [3]Geophysical Fluid Dynamics Institute and the Department of Earth, Ocean and Atmospheric Sciences, Florida State University, Tallahassee, FL, USA, [4]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA, [5]Division of Biostatistics, University of California, San Diego, La Jolla, CA, USA

Abstract Argo-type profiling floats do not receive satellite positioning while under sea ice. Common practice is to approximate unknown positions by linearly interpolating latitude-longitude between known positions before and after ice cover, although it has been suggested that some improvement may be obtained by interpolating along contours of planetary-geostrophic potential vorticity. Profiles with linearly interpolated positions represent 16% of the Southern Ocean Argo data set; consequences arising from this approximation have not been quantified. Using three distinct data sets from the Weddell Gyre—10-day satellite-tracked Argo floats, daily-tracked RAFOS-enabled floats, and a particle release simulation in the Southern Ocean State Estimate—we perform a data withholding experiment to assess position uncertainty in latitude-longitude and potential vorticity coordinates as a function of time since last fix. A spatial correlation analysis using the float data provides temperature and salinity uncertainty estimates as a function of distance error. Combining the spatial correlation scales and the position uncertainty, we estimate uncertainty in temperature and salinity as a function of duration of position loss. Maximum position uncertainty for interpolation during 8 months without position data is 116 $\pm$ 148 km for latitude-longitude and 92 $\pm$ 121 km for potential vorticity coordinates. The estimated maximum uncertainty in local temperature and salinity over the entire 2,000-m profiles during 8 months without position data is 0.66 °C and 0.15 psu in the upper 300 m and 0.16 °C and 0.01 psu below 300 m.

Plain Language Summary Argo-type profiling floats do not receive GPS positioning while under sea ice. Current common practice is to approximate the unknown position by linearly interpolating between the known positions before and after ice cover. This linear interpolation is not the true path that these floats follow with under the ice. What is the uncertainty of this linear approximation? Float position and velocity decorrelate with time—meaning the linear approximation of position tends to be worse as time increases. In our paper, we address the question of measurement uncertainty as a function of time by breaking the problem into two pieces: the position uncertainty as a function of time and the measurement uncertainty as a function of position. Combining these statistics, we estimate uncertainty as a function of time of position loss for temperature and salinity as well as surface fluxes derived from the Southern Ocean State Estimate.

Key Points:
- Argo float position uncertainty when lost under sea ice for 8 months was found to be on the order of 100 km
- Mapping error in salinity and temperature due to 8 months of position loss was found to be 0.15 psu and 0.66 degrees C in the upper ocean
- Mean RMS difference of surface heat and salinity flux between true and interpolated float trajectories was 28.5 W/m^2 and 1.8×10^{-3} kg · m^{-2} · s^{-1}

Correspondence to:
P. M. Chamberlain,
pchamber@ucsd.edu

Citation:
Chamberlain, P. M., Talley, L. D., Mazloff, M. R., Riser, S. C., Speer, K., Gray, A. R., & Schwartzman, A. (2018). Observing the ice-covered Weddell Gyre with profiling floats: Position uncertainties and correlation statistics. *Journal of Geophysical Research: Oceans*, *123*, 8383–8410. https://doi.org/10.1029/2017JC012990

Received 13 APR 2017
Accepted 27 SEP 2018
Accepted article online 2 OCT 2018
Published online 22 NOV 2018

1. Introduction

Antarctic winter sea ice covers an enormous region (18.5×10^6 km^2 average maximum winter extent); 85% of this sea ice is seasonal (Fetterer et al., 2016). The ocean interior circulates heat and freshwater and supports a vigorous seasonal cycle in carbon, oxygen, and nutrients. These processes and the ocean-ice-atmosphere exchanges that occur within the ice-covered region impact the climate both regionally and globally. Formation of Antarctic Bottom Water, one of the densest water masses of the world ocean, occurs in polynyas along the coast of Antarctica in winter; changes in its production, temperature, and salinity over the past 60 years have contributed significantly to changes in the global ocean heat budget and regional freshwater budget (Desbruyères et al., 2016; Purkey & Johnson, 2010, 2012; Rhein et al., 2013). Upwelling of deep water to the surface of the Southern Ocean supplies the source water for formation of both Antarctic Bottom Water and thermocline/Subantarctic Mode Water (SAMW).

Changes in salinity of these source and sink waters have been linked to changes in sea ice formation as well as to ice shelf melt (Haumann et al., 2016; Jones et al., 2015) and are likely associated with changes in the basic freshwater fractionation mechanism associated with sea ice formation near Antarctica and export of that sea ice northward to where it melts, which salinifies the coastal waters and freshens the surface waters of the Antarctic Circumpolar Current (Abernathey et al., 2016). This freshening affects the properties of SAMW when these surface waters are driven northward across the Subantarctic Front by wind (Cerovečki & Mazloff, 2016). Sarmiento et al. (2004) showed that nutrients in the upwelled deep waters that enter the SAMW have significant global impacts on nutrient availability in the upper ocean.

The seasonally ice-covered region of the Southern Ocean is relatively inaccessible to ships during winter. Through growing deployments of autonomous Argo-type profiling floats, many of which now carry biogeochemical sensors (Johnson & Claustre, 2016), these observational gaps are being filled in. However, almost all Argo-type floats are tracked only through satellite positioning at the sea surface. Argo-type floats have a serious vulnerability when operating in the polar oceans: sea ice is very destructive to instruments at and near the surface. To protect floats from sea ice, a temperature-controlled ice avoidance algorithm has been implemented in Argo float buoyancy control software since 2001 (Klatt et al., 2007). This algorithm instructs the float to initiate a descent if the upper ocean temperature is below a threshold that implies the strong probability of sea ice. Float profiles collected under ice are saved in memory for later ice-free transmission without the satellite-derived position fixes of open water Argo profiles. In the Southern Ocean (south of 55° S), there are 11,856 temperature-salinity profiles with position flags indicating interpolation due to lack of position information, from 231 Argo floats since 21 July 2006 (Figures 1 and 2). Most interpolated locations are due to sea ice cover. These data represent 16% of the Southern Ocean Argo data set.

The question we address in this paper is the extent to which analysis and use of these under-ice profiles are compromised because their positions are interpolated. The most common approach to dealing with missing Argo position information is to linearly interpolate latitude-longitude between known positions (Wong & Riser, 2011), although interpolation along contours of planetary-geostrophic potential vorticity (PV) has been proposed as an improvement (Reeve et al., 2015). Linear interpolation is apparent in Figure 1 where the normally curved float trajectories are instead a series of straight lines within the sea ice zone. Clearly, floats do not follow these straight paths, which means, among other things, that their total track length and speed are underestimated (Wong & Riser, 2011). The error and impact of assigning under-ice locations to float profiles that can be tens to hundreds of kilometers from their true position is studied here for the first time; the resulting position errors introduce uncertainties in the attribution of temperature and salinity profile measurements to a given position. We refer to these temperature and salinity profile uncertainties as *representation errors*, which are necessary to characterize float data mapped onto Eulerian coordinates—for instance when assimilated into state estimates (e.g., Mazloff et al., 2010; Nguyen et al., 2017) or used for gridded estimates of transport (e.g., Gray & Riser, 2014). In addition, the lack of position information is not only a concern when using float data in the Eulerian reference frame. In the Lagrangian perspective, interpretations of changes in properties of water parcels often utilize surface observations or products such as air-sea fluxes, surface temperature, ice cover, surface height, and ocean color (D'Ortenzio et al., 2014; Wong & Riser, 2011). The uncertainty in the surface boundary conditions that is introduced by uncertainty in under-ice Argo float position is quantified here.

Some Argo floats are tracked acoustically in the sea-ice zone, between their 10-day satellite position fixes when they profile; such acoustic (or RAFOS) tracking (Rossby et al., 1986) has had widespread success in ice-free areas. Moored acoustic sources have been deployed by the Alfred Wegener Institut (Fahrbach et al., 2011; Klatt et al., 2007) in the Weddell Sea. RAFOS-enabled floats work with a network of sound sources that broadcast sound pulses at specified intervals, permitting tracking if enough sources are within range. Full insonification and acoustic tracking of all autonomous under-ice instruments in the Southern Ocean would be ideal; expense and logistics have made this goal presently impractical. A complication here is that the Sound Fixing and Ranging channel is less efficient at conducting sound travel in the Weddell Sea compared with midlatitudes: this reduces the range of acoustic transmission of polar compared to midlatitude sound sources and therefore requires a higher density of sources (Klatt et al., 2007). Acoustic float tracks are helpful for determining the uncertainty of linearly interpolated track approximations, and the uncertainty in linear approximations is essential in evaluating the scientific effectiveness of acoustic tracking versus interpolation.

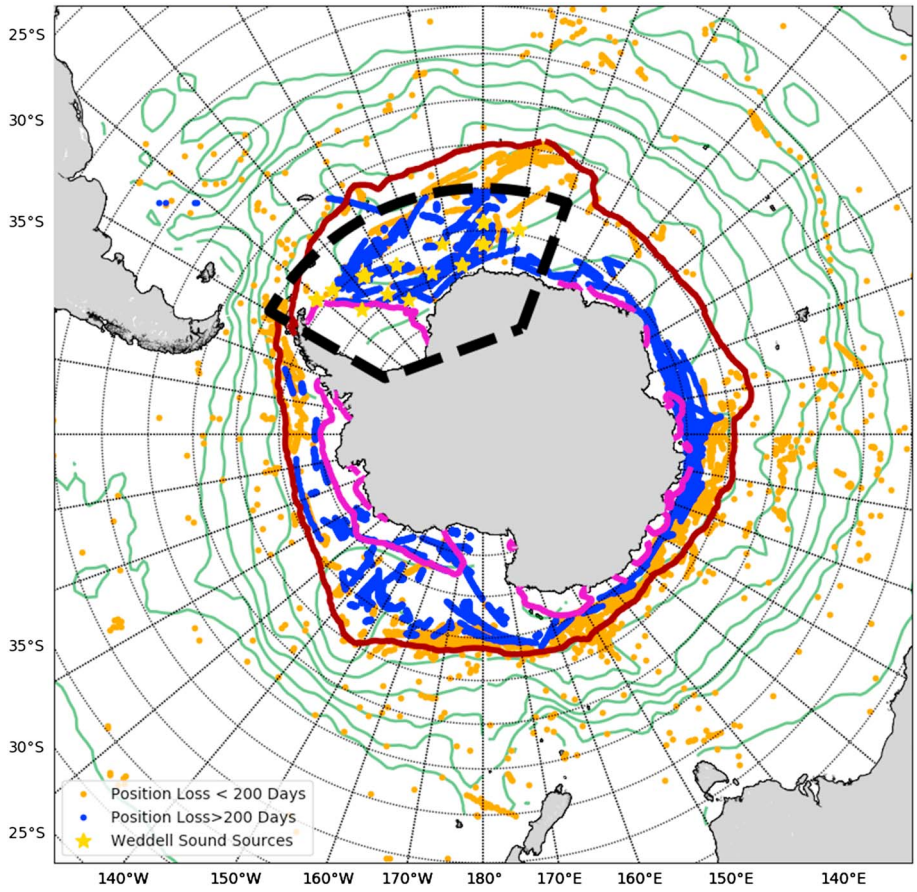

Figure 1. All interpolated Argo float positions in the Southern Ocean. (yellow and blue dots) Interpolated Argo positions with position loss <200 days (yellow) and >200 days (blue). (gold stars) Weddell Sea sound sources (Klatt et al., 2007). (green curves) Average 1,000-m streamlines from Argo data (Gray & Riser, 2014). Black-dashed lines enclose our Weddell Sea study area. Mean September ice edge (red) and mean February ice edge (magenta) from National Snow and Ice Data Center (Fetterer et al., 2016), averaged from 1978 to present.

We present a method to estimate representation error associated with position interpolation of any Lagrangian measurement over any time interval and in any geographic region. We split the problem of finding representation error as a function of duration of position loss (section 4.2.1) into the problem of finding position uncertainty as a function of duration of position loss (section 3) and spatial correlations of observed fields (section 4), focusing on temperature and salinity, and peripherally on uncertainty in air-sea fluxes ascribed to the uncertain profile locations.

In section 3, we estimate the position uncertainty of latitude-longitude and PV coordinate interpolation as a function of duration of position loss by performing a series of data withholding experiments on three distinct data sets: daily RAFOS-enabled float track data, 10-day satellite-tracked float data (non-ice periods only), and model output generated from a particle release simulation (Wang et al., 2014) with a 0.16° Southern Ocean State Estimate (SOSE; Mazloff et al., 2010). We interpolate in both latitude-longitude and PV geometries. The results of the data withholding experiment are used to estimate position uncertainty as a function of time since the most recent position observation for periods of up to 8 months.

In section 4.1.1, we estimate representation error as a function of distance uncertainty by using Ninove et al.'s (2016) approach to compute temperature and salinity correlation within a binned spatial grid in both latitude-longitude and PV coordinates—this gives us correlation as a function of distance. We then use distance as a linear predictor of uncertainty and compute uncertainty in temperature and salinity as a function of distance uncertainty. In section 4.1.2, we reverse this viewpoint and consider the uncertainty in surface fluxes associated with positions of a float lost beneath the ice. We calculate correlation length scales of SOSE heat flux, salinity flux, and buoyancy flux and use these correlations to predict uncertainty in these fluxes as

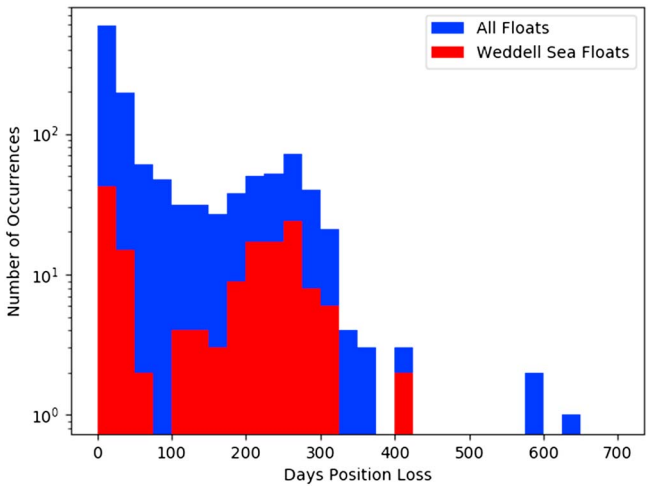

Figure 2. Histogram of total interpolated position time for all Antarctic Argo floats (blue bars, based on all yellow and blue dots in Figure 1) and all Weddell Sea Argo floats (red bars, based on study area in Figure 1).

a function of distance. Finally, in section 4.2.3, we combine these statistics and estimate temperature, salinity, and air-sea flux uncertainty at the interpolated position as a function of position loss time.

To recapitulate, position uncertainty is calculated as a function of position loss time, representation error (mapping uncertainty) is calculated as a function of position uncertainty, and total uncertainty is calculated by combining the two.

2. Float Data and Model Output

The Weddell Sea supports a wind-driven cyclonic gyre; the northern side of the gyre is the southern boundary of the Antarctic Circumpolar Current. For this study, we define the Weddell Gyre as the ocean south of 60° S and between 60° W and 20° E (black-dashed area in Figure 1). The southwest quadrant of the Weddell Gyre is shallower and generally ice covered year round. Sea ice covers the entire Weddell Gyre during austral winter (Figure 1).

The Weddell Sea was chosen for this study because it has a high density of standard satellite-tracked Argo floats compared with other Southern Ocean regions and because of its unique record of under-ice, continuously-tracked, RAFOS-enabled Argo floats deployed through U.S. Argo and AWI (Klatt et al., 2007). The Weddell Sea also has the longest under-ice float records (51 floats since 21 July 2006) of any Southern Ocean region irrespective of insonification.

Depending on their location in the Weddell Gyre, all of which experiences seasonal ice cover, some floats may be forced to operate below ice — without surfacing — for 10 months or more. Figure 2 shows the distribution of periods of position loss for floats circulating in the Weddell Gyre. We observe two major modes of the under ice distribution — transient position loss events lasting 50 days or less and seasonal position loss events lasting between 200 and 300 days. The former represent floats that experience sea ice and then are advected either to an ice-free region or, more likely, out of the gyre entirely. The latter are floats that are caught in the cyclonic Weddell Gyre circulation and experience entire seasons of sea ice cover: these are the floats of primary interest in the uncertainty analysis of this paper.

2.1. Satellite-Tracked Argo Floats

Argo floats typically profile from a maximum depth of 2,000 m to the surface every 10 days, although operational or experimental considerations can change this schedule. While at the surface, satellite positioning information is recorded when possible. All Argo data prior to 17 July 2018, in our specified region, were downloaded from the USGODAE Argo GDAC Data Browser (SEANOE, 2000; regional map of satellite-tracked floats; Figure 3a). An example of these data, from Argo float 5901717, has been plotted in Figure 4, with the under-ice profiles lightened. In general, there are far fewer observations in the southwestern Weddell Gyre, most likely due to a deliberate deployment choice to avoid "losing" floats under multiyear ice (Figure 3).

The Argo user's manual (Carval et al., 2015) lists nine different possible metadata flags for position quality control, although only three are commonly used: flag "1" for good positions, flag "4" for bad positions, and flag "8" for interpolated positions. Argo metadata and flags do not currently include the position interpolation method. Linear interpolation in spatial coordinates is the only method currently used, although adding variants of position reconstruction is under discussion (M. Scanderbeg, personal communication, March 12, 2017).

Only Argo locations with position flag 1 or 8 and Argo temperature and salinity profiles with position flag of 1 were retained for the calculations presented. Of the float data considered, 91 floats were tracked from 10 June 2002 to the present with ARGOS satellites, and 73 floats were tracked with GPS satellites from 12 November 2007 to present. The GPS constellation tracking is more accurate than ARGOS tracking, with stated location uncertainties of better than 7.8 m. The ARGOS constellation uses the Doppler shift of received transmissions to estimate position and is sensitive to the number of received transmissions, and the configuration of the constellation at the time the messages were received. ARGOS positioning has four levels of accuracy ranging from *better than 250-m radius* to *over 1,500-m radius*. For ARGOS satellite-derived positions, 19.4% of posi-

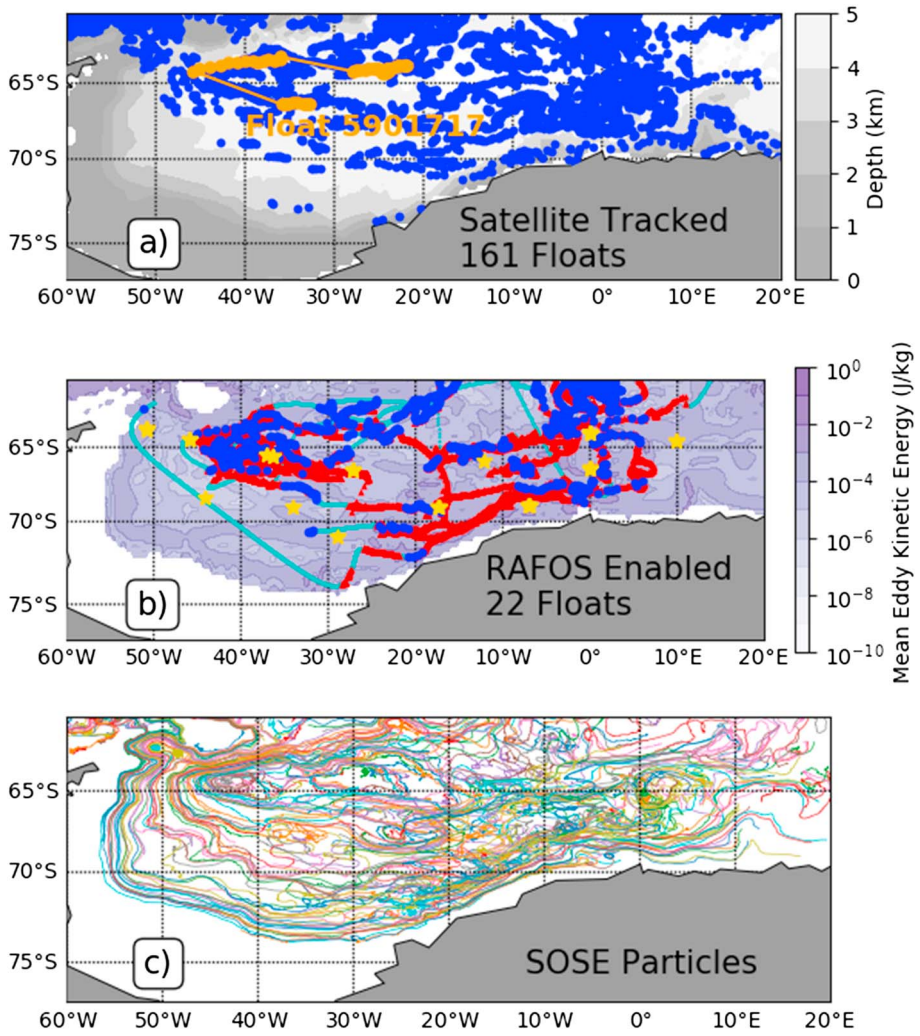

Figure 3. (a) Satellite-tracked Argo float positions (blue dots) in the Weddell Sea with bathymetry (gray shading). Float 5901717 used in Figures 4 and 13 is highlighted in orange. (b) AWI sound sources (gold stars), acoustically tracked positions of RAFOS-enabled Argo floats (red dots), satellite-tracked positions of RAFOS-enabled floats (blue dots), Kalman smoother interpolation (cyan dots), and 950-m mean eddy kinetic energy from SOSE (shading). See Table 1 for list of RAFOS-enabled Argo floats. (c) Two hundred forty SOSE particle tracks of particle release experiment (1,000 particles total) in the Weddell Sea. SOSE = Southern Ocean State Estimate.

tions were deemed *better than 1,500-m radius*, 43.2% were deemed *better than 500-m radius*, and 37.4% were deemed better than 250-m radius. Only three positions were deemed *greater than 1,500-m radius* and were rejected from the data set.

Argo floats that use the ARGOS constellation for positioning and data transmission may spend up to 12 hr at the ocean surface, whereas GPS-tracked floats, which use Iridium satellites for data transmission, typically spend less than an hour. During this period of transmission, Argo floats are essentially undrogued drifters, affected by surface winds and currents. After transmission, floats conduct their profiling mission and are advected by the shear in the ocean velocity field both on descent and ascent. Although techniques have been developed to correct for these dynamical errors (Gille & Romero, 2003), we have not applied them here.

2.2. RAFOS-Enabled Argo Floats

Quantifying position error for under-ice Argo floats, which have linearly interpolated positions between satellite fixes, is greatly facilitated in the Weddell Gyre because of the installed long-term acoustic float tracking system and deployment of RAFOS-enabled (Rossby et al., 1986) Argo floats between 26 February 2008 and 4 February 2013. A subset of 22 of these floats was processed and tracked for this study (Table 1; Boebel, 2009;

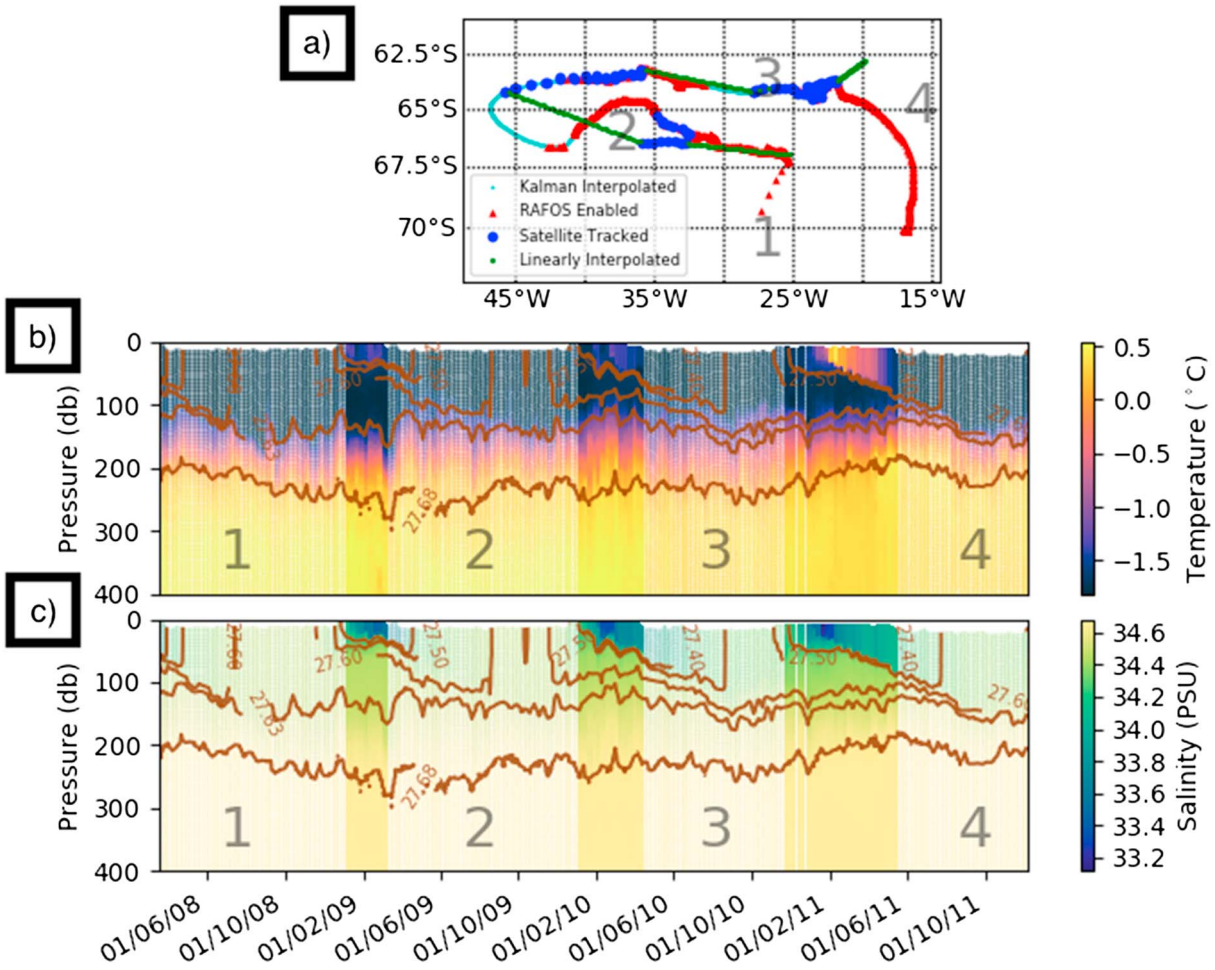

Figure 4. (a) WMO ID Float 5901717 float track (blue dots showing 10-day satellite-tracked positions) in the Weddell Sea. Straight segments (green) are periods of Argo linear interpolation due to position loss in the presence of sea ice. Acoustic tracking has been produced for this float (red; Figure 3b). (b) Potential temperature for WMO ID float 5901717 with superimposed potential density σ_θ (brown curves). (c) Salinity for WMO ID float 5901717 with superimposed potential density σ_θ (brown curves). Numbers and lightened regions in (b) and (c) are periods of position loss due to the presence of sea ice.

Fahrbach & de Baar, 2010). These data spanned 1,806 days and 29,183 position measurements (Figure 3b) and were restricted to the Weddell Gyre.

The 22 RAFOS-enabled Argo floats used here are identical to standard (only satellite tracked) Argo floats, in that every 7–10 days, they collect temperature and salinity profiles and then attempt to surface, transmit their data, and receive a satellite fix. The only difference is that they additionally listen for RAFOS signals while drifting at depth. All Argo float profiles, including RAFOS-enabled float profiles, have been entered into the Argo database as part of the normal Argo data stream. According to Argo database standard practice, under ice profile locations are assigned positions using linear interpolation. (Note that although these 22 floats had

Table 1
RAFOS-Enabled Argo Profiling Floats Tracked in the Weddell Sea Deployed From 26 February 2008 to 4 February 2013

WMO identifiers for RAFOS-enabled Argo floats						
5901716	5901717	5901718	5901720	5901721	5901723	5901724
5901728	5901730	5901731	5901733	5901734	5901735	5901736
5901737	5901738	5901739	5901740	5901741	5901742	5901743
5901744						

Note. Argo data graphics are available at http://www.ifremer.fr/co-argoFloats/.

RAFOS receivers, their Argo database locations are based on satellite positioning and linear interpolation since RAFOS positions were not available; Argo data management does not include RAFOS-tracked positions at this time.)

A network of moored sound sources (Boebel, 2009; Fahrbach & de Baar, 2010) in the Weddell Gyre has been maintained since 1999 by the Alfred Wegener Institute (Klatt et al., 2007; gold stars in Figures 1, 3, 4, and 13). Sound sources broadcast at carefully coordinated daily intervals to avoid interfering with one another. RAFOS-enabled Argo floats listen for acoustic pings and record the time of arrival (TOA) of the sound pulses. By subtracting the known sound source clock time from the TOA, we calculate the time of sound travel. Dividing time of sound travel by the sound speed then gives a range from the float to the sound source. The sound speed is determined from Argo temperature and salinity measurements, which provide an approximate ocean density structure between the float and sound source (McDougall & Barker, 2011). The average sound speed in the upper 1,000 m of the Weddell Gyre was calculated using all Argo temperature and salinity profiles, yielding 1,441.9 m/s. We used this average sound speed for our positioning calculations to account for the complex paths that sound pulses take in the ice-covered Weddell Sea. Float and sound source clock times are subject to errors and offsets, and because of the rapid propagation of sound in the ocean, small timing errors can cause large changes in float track solutions.

The location of a float, which includes positioning error, can be determined if three or more distinct ranges are known. The ARTOA II software package (Wooding et al., 2005) is the most commonly used approach for constructing float track solutions from acoustic data. Typically, the ARTOA II software uses a least squares approach to solve for float position although other schemes are available. Like standard Argo floats, the RAFOS-enabled Argo floats attempt to surface every 10 days and record satellite-derived positions when conditions allow. ARTOA II does not use satellite positioning to constrain float tracks; instead, it informs the user of potential tracking biases by overlaying satellite positioning information onto acoustically derived float tracks and TOA scatter plots.

Unfortunately, tracking floats in the polar oceans is not as simple as in midlatitudes. In midlatitudes, large sound source transmission range (>1,000 km) is achieved through the efficient sound conduction of a broad and deep Sound Fixing and Ranging channel; in the Weddell Gyre, the Sound Fixing and Ranging channel is relatively shallow and narrow. When sound is conducted close to the surface, unknown interactions and reflections with sea ice occur. This causes transmission ranges in the Weddell Gyre to vary from 700 km to as little as 300 km (Klatt et al., 2007). The acoustic tracking data for the 22 floats that we processed from this region frequently had less than three ranges. In addition to this issue, both measurements and sound speed have uncertainties, most notably the slow drift of sound source clocks and float clocks. As will be described in a subsequent publication, the least squares solution commonly used by ARTOA II is susceptible to becoming unstable if float and sound source clock drift has not been precisely accounted for in preprocessing or if only sparse range data is available.

Broader processing challenges can occur when dealing with the specifics of this formulation of acoustic tracking: what quantitatively defensible adjustments to the clock offsets should be made when acoustic ranging differs significantly from satellite positioning? If two acoustic ranges differ significantly, which one should be trusted? Can information learned about the clock drift of one source be propagated to the rest of the float tracking network? Can we estimate float location if we have just one or two acoustic range estimates instead of the three required in a least squares solution?

Motivated by these questions and to overcome the greater challenges of intermittent and noisy range data in the ice-covered regions, we developed a Kalman smoother approach (Rauch et al., 1965) to process these Weddell Gyre RAFOS-enabled floats, which combines intermittent satellite positioning, noisy acoustic TOA, and dynamical information to reconstruct float tracks. The Kalman filter is a minimum mean square error estimator that is commonly used to combine intermittent and disparate sensor observations. In the computationally efficient formulation of the Kalman smoother that we have developed, the state variables and covariances are saved on the forward pass and applied during the backward pass to compute improved state estimates.

Our Kalman smoother method has been designed to solve the entire float system simultaneously—it treats all float positions, float velocities, sound source clock offsets, and sound source clock offset drifts as state variables. The Kalman smoother takes as input all GPS locations and acoustic ranging data. A priori estimates

of state variables are based on crude predictive models

$$x_{t+1} = x_t + v_t \Delta t, \qquad offset_{t+1} = offset_t + drift_t \Delta t \tag{1}$$

where x is the float position, v is the float velocity, *offset* is the sound source clock offset, and *drift* is the sound source clock offset drift. In this formulation, every time a float receives a satellite position on the same day as it acquires acoustic ranging data, the smoother calibrates the sound source clock offset and recalculates the offset drift. Our Kalman smoother approach has been validated using many ARTOA II processed acoustically tracked float trajectories from the Diapycnal and Isopycnal Mixing Experiment in the Southern Ocean experiment, which benefited from both lack of sea ice and sufficient sound sources to allow accurate tracking (LaCasce et al., 2014). For our calculations, we used a 0.05° uncertainty in satellite-derived positions, 10-s uncertainty in acoustic ranging, 3×10^{-3} for process noise in position (units of degrees^2), velocity (units of degrees^2 /day^2), and sound source clock offset (units of seconds^2), and 1×10^{-4} s^2/day^2 in process noise of sound source clock offset drift. The 22 Weddell Gyre RAFOS-enabled floats had satellite positions for 1,155 days on a 10-day sampling interval, acoustically derived positions on 12,281 days on a daily sampling interval, and had no positioning on 10,356 days. The Kalman smoother provides an estimate of float position for every day regardless of how many acoustic ranges are available.

2.3. SOSE Particle Release Simulations

To complement the RAFOS-enabled and satellite-tracked data sets for estimating uncertainties due to position loss while profiling floats are under ice, we conducted particle release experiments in the SOSE (Mazloff et al., 2010). SOSE is an eddy-permitting 0.16° Massachusetts Institute of Technology General Circulation Model, which is fit by constrained least squares to altimetry and water property data sets; depths are discretized into 42 levels of decreasing resolution with a domain spanning from 24.7°S to 78°S. The current SOSE version (iteration 100) spans 6 years (2005–2010). While SOSE has been extensively validated (Abernathey et al., 2016; Cerovečki et al., 2013), SOSE has known errors and biases but is a powerful tool that can be used to diagnose Southern Ocean variability. SOSE currently assimilates all satellite-tracked and linearly interpolated Argo profiles. A module has been developed to simulate and track the release of Lagrangian particles (Wang et al., 2014) at specified depths in the water column. One thousand simulated particles were released at locations distributed throughout our geographic region in water depths greater than 1,000 m. The particles were tracked for 1,500 days starting 1 January 2005, saving the resulting positions every 2 days. All particles were assigned to depth level 23, corresponding to a depth of 950 m. Only particle trajectories within the original region are analyzed here. The final model output contained 611,022 position measurements (a subsample of particle tracks is shown in Figure 3c).

2.4. SOSE Air-Sea Fluxes

One goal of this manuscript is to assess the uncertainty in interpreting under-ice processes that arises from projecting external Eulerian surface boundary conditions onto Lagrangian observations at interpolated positions. Direct flux observations lack the temporal and spatial resolutions in this region to perform this calculation, so we analyzed daily-averaged surface heat flux, salinity flux, and buoyancy flux output from SOSE (Cerovečki et al., 2011). As a validation of the SOSE air-sea fluxes, Cerovečki et al. (2011) show that the air-sea fluxes that are adjusted by SOSE from initial National Center for Environmental Prediction or European Reanalysis-Interim flux fields are more representative of directly observed air-sea fluxes than the initial National Center for Environmental Prediction or European Reanalysis-Interim reanalysis fields.

Following Cerovečki et al. (2011), buoyancy flux is expressed in units of heat-equivalent buoyancy flux and is the sum of air-sea heat flux and freshwater heat-equivalent flux

$$Q_{BF} = Q_{HF} - Q_{FW} = \frac{\rho_0 c_p}{g\alpha} B, \tag{2}$$

where the buoyancy flux B is given by

$$B = B_{HF} + B_{FW} = \frac{g}{\rho_0}\left[\frac{\alpha Q_{HF}}{c_\rho} - \rho_0 \beta S(E - P)\right], \tag{3}$$

and g is the gravitational constant, ρ_0 is a reference density, c_p is the specific heat of seawater, S is the surface salinity, E is evaporation, P is precipitation, and α and β are the thermal expansion and saline contraction coefficients, respectively. In the Southern Ocean, the freshwater contribution to buoyancy flux is significantly

greater than the air-sea heat flux and dominates equation (2) (Cerovečki et al., 2011). For our analysis, we restricted our model output to the Weddell Gyre and to locations where the modeled ice concentration value was greater than 20%. This SOSE output comprised 28,470 daily averaged model values for each flux type.

2.5. Float Velocity Comparisons

Our results rely heavily on the assumption that the position statistics of the RAFOS, satellite-tracked, and SOSE data sets are similar. To validate this hypothesis, we compared the derived speeds of all three data sets. Float speeds are calculated by finding the difference between successive positions and dividing by the positioning time interval. A probability density function of the RAFOS-enabled float (daily sampling), satellite-tracked float (10-day sampling), and SOSE particle (2-day sampling) speeds is shown in Figure 5. As can be seen by the progressive leftward shift in the peak of the normalized histogram, the mean speed of the more frequently tracked RAFOS-enabled float data set is 2.75 km/day, the mean speed of the satellite-tracked float data set is 2.16 km/day, and the mean speed of the SOSE particle experiment is 1.90 km/day.

There is no difference between the pressure case of RAFOS-enabled and satellite-tracked Argo floats, so we expect their motion through the water to be the same. To further investigate the 0.6-km/day difference between the RAFOS- and satellite-tracked mean speeds, we subsampled the RAFOS-enabled float positions at a 10-day interval — the same used by satellite-tracked floats; the position subsampled RAFOS-enabled data set had a mean speed of 2.32 km/day. In addition to the tracking method, the main differences between the subsampled RAFOS-enabled data set and the satellite-tracked data set are that the former contains far fewer floats and also samples in the southern (and ice covered) regime of the Weddell Gyre. The subsampled RAFOS-enabled data set consisted of 2,239 calculated speed instances compared to 3,640 speed instances in the satellite-tracked data set. We suspect that the difference of 0.16 km/day is likely due to statistical uncertainty due to sample size and comparison of the derived speeds (and positions) of these data sets is appropriate, although it is possible that either there exists a Weddell Gyre circulation regime sampling bias in these data sets or a tendency of the Kalman smoother to overestimate velocities.

As shown in the previous calculation, when float tracks are linearly interpolated, calculated speed is biased low because the distance between linearly interpolated positions is shorter than the true float path. To further compare the 10-day sampling interval of satellite-tracked floats and the daily sampling interval of RAFOS float tracking, we calculated the difference in mean speed between acoustically positioned float tracks and the interpolated float tracks of the RAFOS-enabled float data set during all periods of position loss (Figure 5b). These data are color coded by duration of position loss. In general, we see greater speed differences during periods of longer position loss, as expected.

There are two limits for the speed difference between the acoustically positioned float tracks and linearly interpolated float tracks: the turbulent limit where the float speed difference is equal to the acoustically tracked speed and the mean limit where the float speed difference is equal to 0. We observe a significant number of float paths approaching both of these limits. Average speed differences between RAFOS-enabled float velocities and interpolated Argo float velocities were 1.59 cm/s for periods of position loss greater than 180 days and 0.39 cm/s for periods of position loss less than 180 days. This discrepancy in speed underscores the need for an assessment of the uncertainty of the linear interpolation approximation. Wong and Riser (2011), analyzing seasonal ice zone Argo floats in the East Antarctic sector, similarly noted that the 10-day satellite derived tracks of ice-free Argo floats are almost 3 times longer than when those are linearly interpolated through the sea ice season.

3. Position Uncertainty Estimate

Autonomous floats are increasingly being used for under ice observations, for example, 30 SOCCOM floats are under ice in austral winter 2018. Linear interpolation of latitude-longitude is the commonly used (and accepted) approximation for position while under ice, but uncertainties arising from this approximation have never been quantified. A related question is whether linear interpolation in latitude-longitude is the most reasonable choice or might another interpolation scheme be better? Bathymetry has been demonstrated to

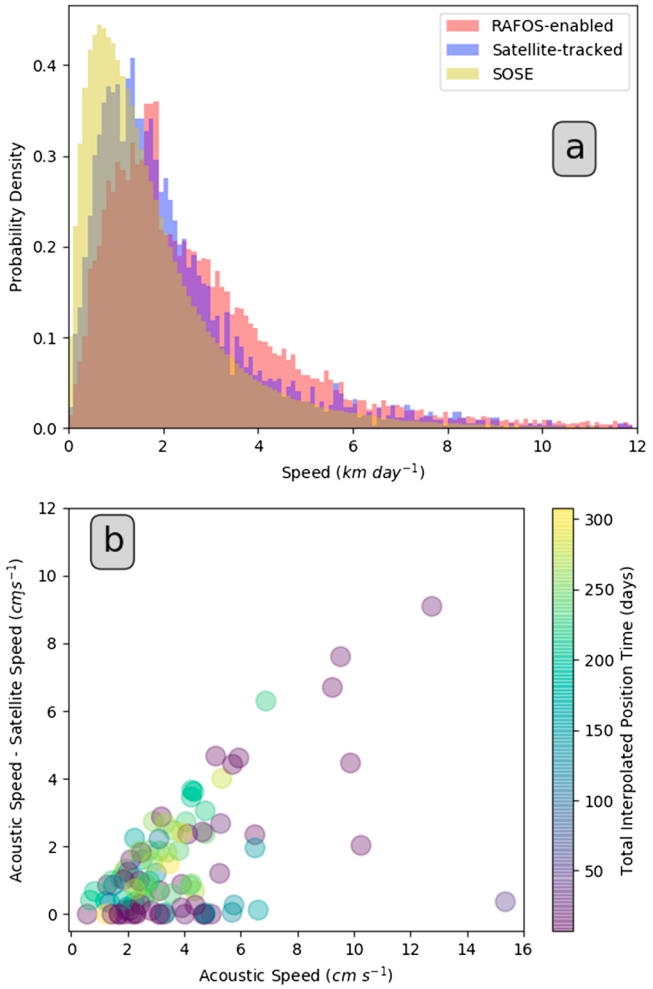

Figure 5. (a) Normalized probability density function of RAFOS-enabled float speeds (red), satellite-tracked float speeds (blue), and SOSE particle speeds (yellow) in the Weddell Sea; (b) mean acoustically derived speed minus mean interpolated speed for all periods of total interpolated position time in the RAFOS-enabled float data set. Color of points indicates duration of position loss. SOSE = Southern Ocean State Estimate.

affect float trajectories in the Southern Ocean, for example, floats in equivalent barotropic flows tend to follow PV contours (LaCasce & Speer, 1999; Reeve et al., 2015), and isolated bathymetric features form Taylor columns that steer float trajectories (Meredith et al., 2015); these examples suggest that a PV-contour following approach might better approximate the trajectories and improve the uncertainties.

We use float and model position data sets, including highly resolved RAFOS-tracked positions as well as 10-day GPS-tracked profiling float positions, to estimate position uncertainty as a function of time of position loss. We estimate position uncertainty by withholding known position information and linearly interpolating in both latitude-longitude and PV coordinates. These linearly interpolated tracks are then compared to known position information, and statistics are averaged.

3.1. Methods

We assess the uncertainty in position loss due to lack of tracking with trajectories for our three independent data sets during periods of complete tracking without sea ice cover (satellite-tracked, RAFOS-enabled, and SOSE particle release) and compare the actual positions with interpolated positions during simulated sea ice cover. Ice cover is simulated by the application of a data withholding experiment; as the name implies, the experiment withholds information from a test case and compares the conclusions drawn from the limited information test to those of the full information case. In this application, we compare a specified duration of float or particle track locations (which we call a *record*) to the positions generated by a linear interpolation of the endpoints of this record (see Figure 7a). The time of each position as well as the great circle distance between the actual and linearly interpolated positions are saved.

All interpolation in this publication is linear and is carried out in either latitude/longitude coordinates, which is the method applied in the Argo float data set, or PV coordinates. The location of any linearly interpolated profile is calculated in each axis by multiplying the axis distance between record end points by the ratio of time of the interpolated profile to total time of the record. Interpolation in latitude and longitude coordinates is straightforward. It is important to note that linear interpolation in latitude-longitude does not necessarily follow a great circle path.

Linear interpolation in PV coordinates requires more work. PV conservation in barotropic planetary-geostrophic flows is expressed as

$$\frac{D(PV)}{Dt} = 0, \qquad PV = \frac{f+\zeta}{h} \approx \frac{f}{h}, \tag{4}$$

where PV is potential vorticity, f is the Coriolis parameter, h is the water depth, and ζ is the relative vorticity. Under these conditions, PV is a material invariant and barotropic flows align with fh^{-1} contours. At high latitudes, such as in the Weddell Sea, where the water column is much less stratified that at lower latitudes, flow is close to equivalent barotropic (unidirectional from top to bottom, although vertically sheared) and thus likely to follow fh^{-1} contours. Consequently, in PV coordinates, our two axes of consideration are along fh^{-1} contours and across fh^{-1} contours. PV at the beginning and end point of a record is not usually the same (hence floats are not perfectly conserving PV). We linearly interpolate along the two axes: along-PV and across-PV.

PV (fh^{-1}) is computed using the Gaussian smoothed 0.16° bathymetry from SOSE, which is based on the Earth Topography Five-Minute Grid and Smith and Sandwell; these data sets are binned to a 0.25° resolution and then interpolated to the SOSE grid. A steepest descent (or ascent depending on the change in PV) algorithm is applied to the location at the end of the record to find the across-PV axis, and an along-gradient algorithm is used from the start of the record to find the along-PV axis (examples in Figures 6a–6c). The intersection of the along-gradient and steepest descent algorithms define the lengths of the PV axes. The two axes are

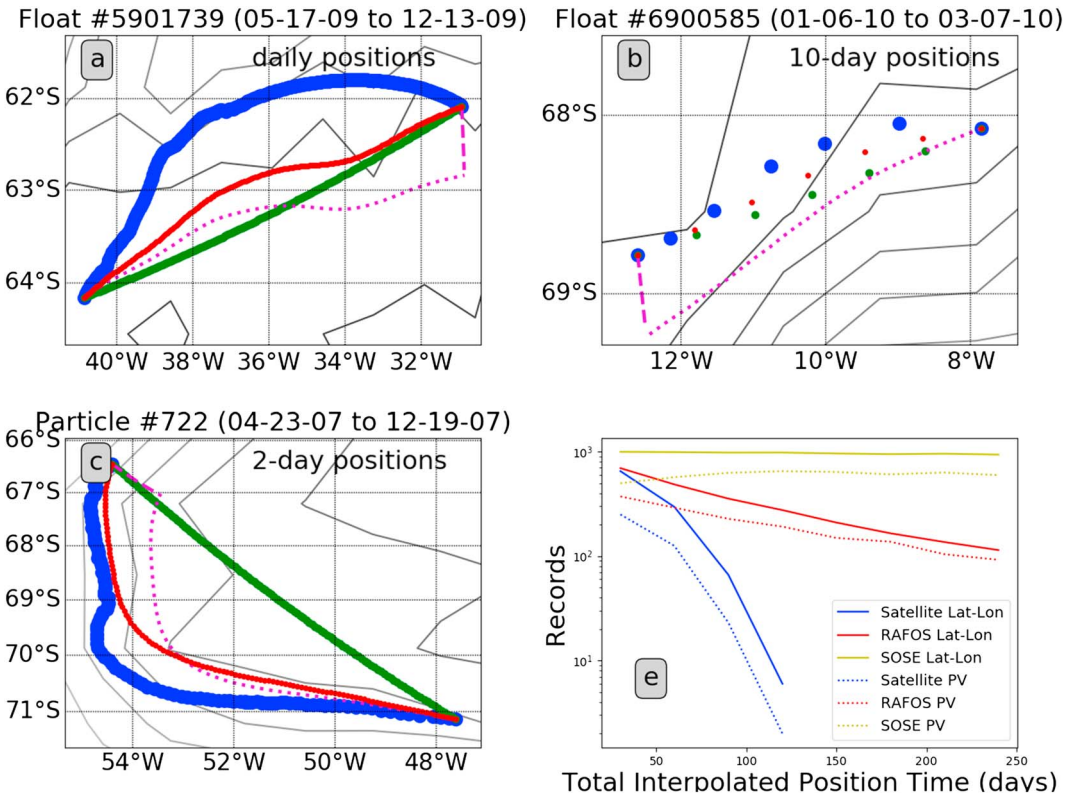

Figure 6. Selected float track data for RAFOS, satellite, and SOSE particle data sets. Actual trajectories (blue dots) are compared with PV interpolation (red dots) and linear interpolation (green dots). Along PV axis (pink dots), across PV axis (pink dashes), and bathymetry (black lines) are provided for context. (a) RAFOS tracking of of WMO float ID 5901739 for 210 days from 17 May 2009 to 13 December 2009; (b) satellite tracking of WMO float ID 6900585 for 60 days from 6 January 2010 to 7 March 2010; (c) SOSE particle 722 subsampled for 240 days from 23 April 2007 to 19 December 2007; (d) size of all data sets: satellite (yellow), RAFOS (red), and SOSE (blue) for linear (solid) and PV (dots) interpolation by time interval. SOSE = Southern Ocean State Estimate; PV = potential vorticity.

then linearly interpolated to find the track in PV coordinates. Bathymetry is, in general, nonlinear; the track generated by reversing the start and end point may not be the same as the original. We calculate both the forward and backward PV interpolations and save the record with the minimum mean error. Records with no intersection between the along- and across-PV contours were rejected from our analysis.

The character of position uncertainty as a function of time depends on the total time of position loss (record length, i.e., the total amount of time under ice), and the length of time since the position was measured for that given record length. We define these two separate quantities as *total interpolated position time* (TIPT) and *interpolated position time* (IPT). TIPT is identical to record length. We have used TIPTs of 1 to 8 months (30 to 240 days), in increments of 1 month. IPT ranges from 0 to TIPT, in increments of 1 day. To estimate position uncertainty as a function of TIPT and IPT, we use float position records that are complete (no missing positions) and perform data withholding experiments. We randomly sampled 1,000 records for each specified period of TIPT and calculated the ensemble standard deviation relative to linear interpolation for IPT intervals of 1 day within the TIPT. This procedure was then repeated for a range of TIPTs, from 1 to 8 months in monthly intervals. To estimate a robust standard deviation of the position uncertainty for a given IPT (within a given TIPT), we repeated this data withholding experiment many times using many different observed records.

To ensure the position data were appropriate for estimating the impact of linear interpolation in the Weddell Gyre, we imposed several criteria: records with data gaps 3 times longer than the sampling interval (6 days of position loss for SOSE, 30 days of position loss for satellite tracking, and 3 days of position loss for RAFOS-enabled tracking) were rejected. Similarly, records with positions outside of the Weddell Gyre were rejected. Lastly, in order to enforce a degree of independence, we allowed no greater than a 50% time overlap in records.

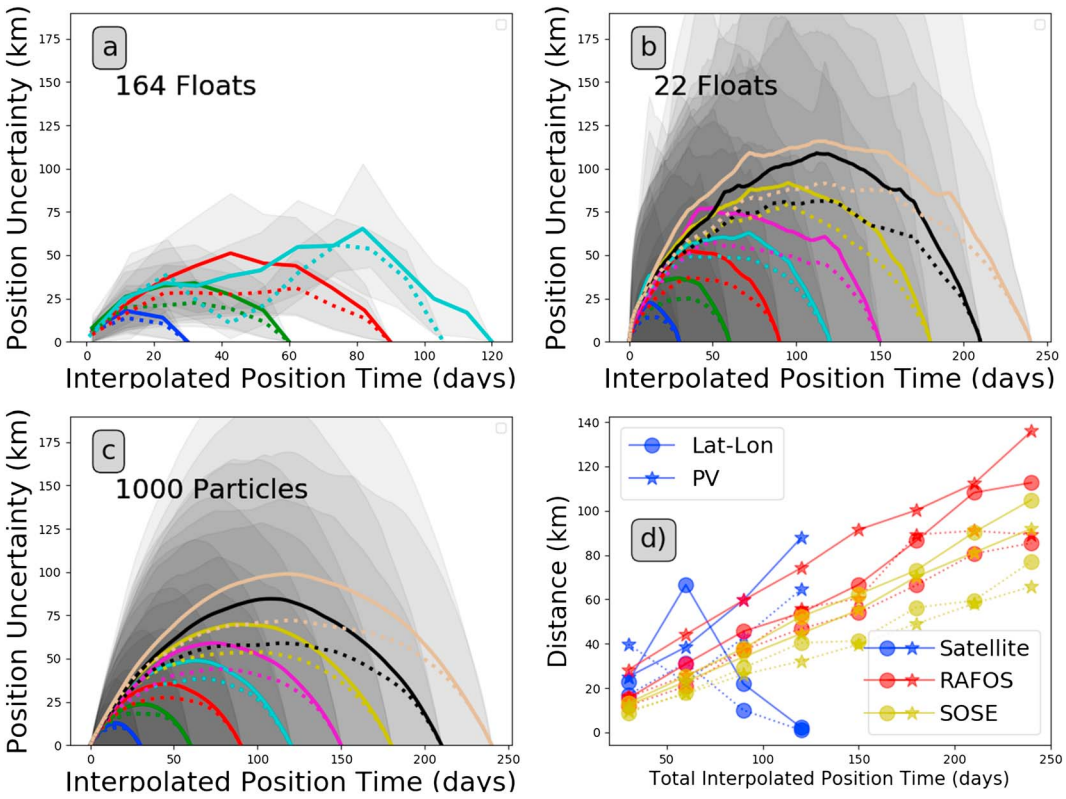

Figure 7. Position uncertainty as a function of interpolated position time, based on the difference between the actual track and the linear and potential vorticity interpolations of the track: (a) satellite positioning derived standard deviation of position uncertainty from 164 floats; (b) RAFOS positioning derived standard deviation of position uncertainty from 22 floats; (c) SOSE-derived standard deviation of position uncertainty from 1,000 particles. For all uncertainty estimates (a–c), colored curves are mean position uncertainty of linear (solid) and potential vorticity (dots) interpolation in monthly time intervals (total interpolated position time). Gray-shaded gray areas are standard deviation in the uncertainty estimate. (d) Maximum seasonal position uncertainty estimates for all three data sets for summer (stars) and winter (circles) with linear (solid) and potential vorticity (dots) interpolation. SOSE = Southern Ocean State Estimate.

The entire Weddell Gyre is ice-covered during austral winter. Consequently, the number of satellite-positioned float tracks that satisfied our linear interpolation criteria decays significantly in winter (data density in Figure 6d). Because of this, we were only able to compute statistically significant satellite-tracked position uncertainty for 4 months. This was not an issue for the acoustically tracked floats and SOSE-simulated floats, for which under-ice positions are known. Also, because of the relatively long sampling interval (10 days) of the Argo float tracks, we binned and averaged the mean satellite-tracked data set position uncertainty estimates in 10-day segments. This was not necessary for the acoustically tracked floats and the SOSE output due to the much higher frequency of positioning.

In the Southern Ocean (Figure 1), the sea ice edge moves from north to south from winter to summer. In the Weddell Gyre, the longest interpolation-free records are located in the northeast. The northeastern region is dynamically different from the west, as seen in the kinetic energy field (regional map in Figure 3b). This may cause a sampling bias in the satellite-tracked uncertainty estimate, which is based on records with no missing positions, as these ice-free records are biased toward the northeastern Weddell Gyre. Our estimate in latitude-longitude assumes position uncertainty to be isotropic within this region, whereas the estimate in PV coordinates accounts for bathymetrically controlled fronts and current meanders. Both estimates assume the position uncertainty statistics are unvarying in time; in the following section, we show a seasonal analysis that supports this assumption.

3.2. Results

The standard deviation of the position uncertainty relative to linear interpolation as a function of IPT, for all TIPTs, data type, and interpolation scheme, is shown in Figure 7. The interpolated latitude-longitude position

uncertainty estimates derived from the satellite- and the RAFOS-tracked data sets show excellent agreement for all TIPTs for which we had sufficient data to make these estimates. The SOSE position uncertainties are consistently lower than those from both float analyses. Comparisons of PV interpolation to linear interpolation were spatially structured and generally confirmed the hypothesis that floats follow PV contours: PV interpolation was better than linear interpolation for all data sets.

The averaged position uncertainty curves (Figures 7a – 7c) have a quadratic shape as a function of IPT, for all three data sets, with highest position uncertainty at the middle of the time interval and with that maximum uncertainty increasing with total period of position loss. This maximum amplitude of position uncertainty increases linearly with period of position loss (Figure 7d). The quadratic shape is clearly defined for the SOSE model (2-day timing) and similarly for the relatively continuously tracked RAFOS data (1-day tracking). The satellite-tracked results (10-day interval) are much more variable especially for the longer position loss periods (3 and 4 months), likely due to the smaller number of long records available for analysis since they must be collected during ice-free periods (Figure 7a).

The position uncertainty curves are summarized by selecting the maximum position uncertainty for each monthly time period (TIPT), plotted in Figure 7d. The maximum value for a TIPT of 8 months of position loss, from the interpolated latitude-longitude RAFOS-tracked data set (the value we report in our summary and abstract), is 116 ± 148 km and increased essentially linearly from 0 TIPT at a rate of 0.48 km/day. PV interpolation for the same TIPT and data set was 92 ± 121 km and increased 0.38 km/day. The satellite-tracked data set latitude-longitude interpolation position uncertainty increased 0.54 km/day. PV interpolation position uncertainty for the same data set increased at 0.47 km/day. The high 4-month PV interpolation position uncertainty estimate is an unexpected result and could be an outlier because it does not match the linear trend of the rest of the graph and because of the small number of 4-month satellite-tracked records that met our PV interpolation criteria. The SOSE particle release simulations have a categorically smaller maximum position uncertainty. Latitude-longitude and PV interpolations yield a position uncertainty of 98 ± 130 and 72 ± 97 km for 8 months of position loss, respectively. Position uncertainty of the SOSE data set increased at a rate of 0.41 km/day for linear interpolation and 0.30 km/day for PV interpolation. The ranges of these position uncertainty estimates (Figure 7d) suggest that SOSE underestimates the position uncertainty compared with both the satellite- and RAFOS-tracked data sets by a factor of 80%. Note that the position uncertainty estimates of each of the three data sets are within a standard deviation of the other two, which means that the differences might not be significant.

Figure 7d also shows the seasonal differences in maximum position uncertainty. Summer RAFOS-derived position uncertainty is generally higher than winter position uncertainty; conversely, winter SOSE-derived position uncertainty is categorically higher than summer position uncertainty. Satellite-derived position uncertainty (a more variable signal) is higher in summer or winter depending on the prescribed TIPT. Maximum difference between the winter- and summer-estimated positions uncertainties is for the 120-day satellite-estimated position uncertainty, which seems to be an outlier. All other winter-summer position uncertainty differences were within 15%. From this, we conclude that there is an insignificant seasonal difference in the position uncertainty statistics, which suggests similarities in the under-ice and open ocean eddy fields.

The spatial structure of linear interpolation minus PV interpolation for the SOSE data set is shown in Figure 8 and adds some qualification to the hypothesis of Reeve et al. (2015) that floats follow PV contours in the Weddell Gyre. PV interpolation is generally more accurate at the edge of the gyre in regions of consistent and steep bathymetry and fails in relatively flat areas or regions of complicated bathymetry. This analysis, although intuitive, may be biased by SOSE's underestimation of position uncertainty.

The quadratic shape of the position uncertainty curves is discussed from basic principles in the next subsection. The equivalence between the satelite- and RAFOS-tracked results and difference from the SOSE results are also discussed, with suggestions for estimating position uncertainty in other ice-covered regions that do not have acoustic tracking.

3.3. Discussion

3.3.1. Quadratic Form of Position Uncertainty

The mean position uncertainty curves for each TIPT shown in Figures 7a – 7c have distinctive quadratic shapes, especially when many records are available. Float position decorrelates in time (Balwada et al., 2016). When a float is under ice, we have in hindsight the most accurate position information about the float both immedi-

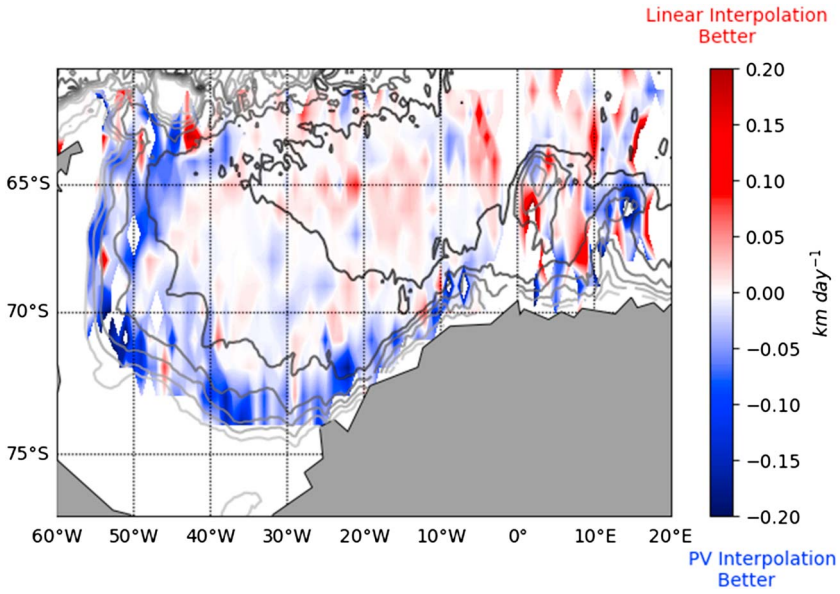

Figure 8. Spatial distribution of linear interpolation uncertainty minus PV interpolation uncertainty (km/day) of Southern Ocean State Estimate data withholding experiment. Red and blue coloring represents the spatially binned and averaged maximum position uncertainty for linear and PV interpolation normalized by period of position loss. Gray lines represent isobaths in 1-km intervals. PV = potential vorticity.

ately after the float submerges under the ice and immediately before the float emerges from the ice. Averaged over many realizations, we observe that float position uncertainty is maximum when IPT is half of TIPT. This is clearly because this time is most removed from the start and end times, when the position is known. Less obvious is why the functional relationship of position uncertainty to time is quadratic.

A well-developed body of literature surrounding the so-called *Brownian Bridge* explains continuous-time stochastic processes with fixed probabilities at both the start and end points. As we will show below, this mathematical construction describes the quadratic shape of the uncertainty curves shown in the previous section and provides some insight into the oceanic mechanisms at play. We will first show that our interpolation process matches the definitions of a Brownian Bridge in one dimension, then expand to show the Brownian Bridge criteria are satisfied in two dimensions, and finally introduce advection.

As the simplest example of a Wiener process, let $x(t) \in \mathbb{R}$ be Brownian motion observed on $0 < t < T$: that is, to say that the motion of the float path is random. By definition for Brownian motion, the expected value $E[x(t)] = 0$, or the motion is unbiased, and the covariance is $Cov[x(t_1), x(t_2)] = \sigma^2 t_1$, where $t_1 < t_2$for some $\sigma > 0$. In the context of the ocean's velocity field, σ^2 has units of diffusivity and is the instantaneous variance at any time t of float velocity perturbations, which are assumed to be uncorrelated over time. This variance results from any random process that affects float position—this could be shear in the upper ocean, advection from wind before or after positioning, perturbation from the eddy field, and so forth. It is important to note that σ^2 is not diffusivity as typically defined (LaCasce et al., 2014) because it is calculated using the absolute distance away from an interpolated position track instead of the actual mean flow.

Suppose the process is not observed over an interval $0 < t < T$. Its interpolated positions, as described in section 3.1, are

$$x_L(t) = x(0) + \frac{t}{T}[x(T) - x(0)]. \tag{5}$$

Conditioned on the values $x(0)$ and $x(T)$, the anchored process $x(t)|x(0), x(T)$ is a Brownian bridge (Revuz & Yor, 2013). Then the following are well known:

$$E[x(t)|x(0), x(T)] = x_L(t), \qquad Cov[x(t_1), x(t_2)|x(0), x(T)] = \sigma^2 \frac{t_2(T - t_1)}{T}.$$

In particular, the mean squared error is

$$E[x(t) - x_L(t)|x(0), x(T)]^2 = Var[x(t)|x(0), x(T)] = \sigma^2 \frac{t(T - t)}{T}.$$

Note that this expression does not depend on $x(0)$ and $x(T)$, so it also holds *unconditionally* at $x(0)$ and $x(T)$:

$$E[x(t) - x_L(t)]^2 = \sigma^2 \frac{t(T-t)}{T}.$$

$$\sqrt{E[x(t) - x_L(t)]^2} = \sqrt{\sigma^2 \frac{t(T-t)}{T}}. \tag{6}$$

This is similar to the quadratic structure that we see in our results.

To show that this is a robust result in two dimensions, let $(x(t), y(t)) \in \mathbb{R}^2$ be composed of two independent Brownian motion processes $x(t)$ and $y(t)$ with the same variance. Let d denote Euclidean distance. Based on the results of the previous section,

$$E\left\{d^2\left[(x(t), y(t)), (x_L(t), y_L(t))\right]\right\} = E[x(t) - x_L(t)]^2 + E[y(t) - y_L(t)]^2 = 2\sigma^2 \frac{t(T-t)}{T}, \tag{7}$$

which is of the same functional form as the one-dimensional case and matches the quadratic structure seen in our results (Figure 7).

Finally, consider the addition of advection to our two-dimensional diffusive process: A simple advective-diffusion process with a constant velocity field (u, v) can be described as $(x(t) + ut, y(t) + vt)$, where $x(t)$ and $y(t)$ are independent Brownian motion processes as before. Conditioned on two known locations $(x(0), y(0))$ and $(x(T) + uT, y(T) + vT)$ at times $t = 0$ and $t = T$, the process in between is a Brownian bridge with mean

$$\left(x_L(t) + ut, y_L(t) + vt\right).$$

This is also the same as linear interpolation between the two points. Subtracting the mean gives the Brownian bridge process $(x(t), y(t))$ in $\mathbb{R}^2$, which is the same as in the previous subsection. Therefore, the position uncertainty that we have estimated from observations should be approximated by (7). This implicitly assumes that over many realizations of the data, tracks generated from linear interpolation approximate the mean circulation, and any deviations from the latitude-longitude or PV interpolation are due to float velocity perturbations caused by the mesoscale eddy field: random processes explain the growth of position uncertainty after a float is lost under the ice and the decline of position uncertainty before a float emerges from the ice. Over many realizations, the float velocity perturbation variance (σ^2), which affects the growth and decline of position uncertainty, should be the same. The Gaussian statistics that this method assumes might break down in fronts or regions of strong velocity gradients. In those regions, other methods of position uncertainty analysis should be considered.

The float velocity perturbation variance σ^2 estimated by fitting the position uncertainty curves in Figures 7a–7c ranges from 159 to 2,614 m^2/s. It is smallest for shortest TIPT and increases with TIPT. If σ^2 represented the actual lateral diffusivity, it would be independent of TIPT. Its dependence on TIPT results from calculating σ^2 relative to the linearly interpolated tracks rather than the actual (unknown) mean flow. That is, unsurprisingly, the longer the period of linear interpolation, the less representative the linearly interpolated track is of the actual mean flow.

3.3.2. Application to Ice-Covered Regions Other Than the Weddell Gyre

Can we usefully apply this method of deriving position uncertainty statistics in other ice-covered regions where RAFOS-tracked floats are not available? The agreement between the linearly interpolated satellite- and RAFOS-tracked results (Figure 7) suggests that we can, recognizing the sampling interval (10 days) for the satellite-tracked floats is longer than for RAFOS tracking (1 day) and there are shorter and fewer continuously tracked records (during ice-free periods) for satellite-tracked floats. The detailed position uncertainty curves for these patchier satellite-tracked records have a more variable mean and standard deviation because of the lack of samples, but the maximum uncertainties appear robust in this Weddell Gyre example. This may seem surprising given that Figure 5a shows that the frequency of position sampling matters in the composition of the velocity spectrum. A wide range of physical phenomena affects float position on different space and time scales, from mesoscale variations and basin circulations to eddy instabilities and inertial motion. Ten-day sampling may underresolve this mesoscale variability, and because of this, the satellite tracking may be expected to underestimate some of the true position uncertainty.

As seen in Figure 7d, these potential sources of difference between the satellite-tracked and RAFOS-enabled data sets do not seem to matter. We conclude that even though the RAFOS-tracked floats have a higher mean

daily velocity and greater probability of higher speeds (section 2.5; Figure 5a), the higher frequency components of the velocity spectrum cancel out on time scales of months, so that over many realizations, the position uncertainty estimates are the same. Therefore, we suggest that position uncertainties using satellite-tracked floats alone could be informative, given that RAFOS tracking is not available in any other Southern Ocean ice-covered regions.

The difference between the float and the 0.16° SOSE results suggests caution in using SOSE particle tracking to estimate position loss uncertainty, although the quadratic functional form is correct. SOSE is an eddy-permitting simulation; hence, the model resolution of 0.16° does not capture the first baroclinic Rossby radius at these higher latitudes, which may underestimate the uncertainty due to an unresolved mesoscale field. Additionally, ice cover may mask the dynamic sea surface height derived from satellite products and used by SOSE. With few advective measurements in the Weddell Gyre, SOSE may underestimate the true mesoscale variability. We see this in the results as SOSE position uncertainty estimates are universally lower than either the satellite positioning or RAFOS-enabled track estimates (Figure 7d). While SOSE is quantitatively different from in situ measurements, the linear increase of the SOSE uncertainty estimate is qualitatively the same. This suggests that a higher-resolution SOSE model with a more energetic mesoscale eddy field might more accurately reproduce position uncertainty statistics.

With planned missions to deploy floats in the inflow of the southwestern sector of the Weddell Gyre, we anticipate that floats under ice for two ice seasons will become more common. It is tempting to extrapolate the straight lines of Figure 7f to estimate position uncertainty of floats with position loss greater than 8 months. It is not clear, though, that the uncertainty will remain linear for periods longer than this.

4. Temperature, Salinity, and Air-Sea Flux Spatial Correlation and Uncertainty

Temperature and salinity are spatially and temporally correlated. Many types of surface processes—solar forcing, buoyancy or heat fluxes from storms, and sea ice melt—span large spatial areas. Our overall goal is to estimate the representation error that the linear interpolation of position introduces in temperature and salinity products. Representation error is independent of float sensor error. Applications like state estimation or objective mapping incorporate observations and require a robust and accurate accounting of all uncertainties in the observations, which include both representation and sensor error.

In the previous section, we estimated how far floats might deviate from interpolated latitude-longitude or PV coordinates; this is the distance for which we will calculate the strength of the oceanic connection. In this section we calculate the correlation length scales of temperature, salinity, and associated fluxes; this is the strength of the oceanic connection at all distances. We then combine correlation length scales and position uncertainty to estimate the representation error.

4.1. Methods

The method for calculating temperature and salinity spatial correlations from discrete, randomly spaced Argo floats that do not measure the entire field continuously is described in section 4.1.1 and is based on Ninove et al. (2016). For analysis of Lagrangian data following a float, gridded Eulerian fields are often employed. Float position uncertainty introduces representation errors when using Eulerian fields; the representation errors are estimated from spatial correlations. In section 4.1.2, we estimate spatial correlation scales of temperature and salinity surface flux from gridded SOSE output. Finally, in section 4.1.3, we use distance as a linear predictor to estimate the representation error field of using one temperature and salinity profile to approximate another at a specified distance.

4.1.1. Temperature and Salinity Correlation Length Scales

Resolving the statistics of a field with irregular space and time sampling is a common problem for oceanographers. We use a modified version of the approach in Ninove et al. (2016), who assembled float data collected in the same time frame, detrended the data, and binned the data by distance to create a discrete- and time-dependent data set from which they estimated a continuous time independent correlation field as a function of distance. In particular, they calculated the seasonally corrected anomaly field by first subtracting the 2009 World Ocean Atlas 1° objectively mapped annual mean climatology (Antonov et al., 2010; Locarnini et al., 2010) and then removing a localized seasonal climatology calculated from in situ observations.

Our goal is to assess the induced representation error of projecting a true profile location onto a linearly interpolated position. Observations made in the same season experience similar forcing with relatively little

spatial variability over the scale of the Weddell Gyre, which increases their correlation (reduces representation error). Leaving the seasonal signal of the temperature and salinity profiles intact reflects knowledge that the true observation and the linearly interpolated observation are recorded at the same time. Consequently, we move the profile only in space and not time, so we depart from Ninove et al. (2016) and remove only the SOSE temperature and salinity mean at the location of the profile.

Most Argo floats profile at 10-day intervals, so, to prevent successive measurements from the same float being compared against itself, we bin the data set by weekly intervals. Next, we calculate the distance between every float pair in a given week. Paired measurements for all depths and densities are binned in increments of 35 km. Ninove et al. (2016) chose to bin results by their zonal and meridional distance. It is hypothesized that water properties should have higher correlation along streamlines. LaCasce and Speer (1999) found that modeled particles were 6 times more dispersive along PV coordinates as across. Reeve et al. (2015) objectively mapped the subsurface temperature maximum using similar reasoning but with length scales determined by the realities of data density. Building on this, it is more useful to bin temperature and salinity data by absolute distance in latitude-longitude and PV coordinates than by the axis of each metric. A description of our method to calculate PV coordinates can be found in section 3.1.

We calculate the correlations between measurements in these distance and depth bins using a Pearson r test masked at a 90% confidence level using a temporal decorrelation of 35 days, determined by Gille and Kelly (1996) using altimetry data averaged over the whole Southern Ocean.

Correlation is expressed as

$$\rho = corr(X, Y) = \frac{cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}. \tag{8}$$

In words, this is an expression for how X and Y change together scaled by the standard deviation of both X and Y, where X and Y represent the binned pairs of temperature or salinity, σ is the standard deviation, and μ is the mean.

The vertical sampling intervals reported for individual float profiles are not uniform. We linearly interpolate all profiles to World Ocean Atlas 2013 version 2 depths (Locarnini et al., 2013; Zweng et al., 2013), to the maximum depth of 2,000 m for Argo floats. This is sufficient to resolve significant features such as the mixed layer and pycnocline. Additionally, we interpolate these data to prescribed levels of potential density referenced to surface pressure using the Gibbs Sea Water package (McDougall & Barker, 2011) in 0.05-kg/m^3 increments. Observed densities varied from 1,026.2 to 1,028.2 kg/m^3, although there were only enough data to calculate correlations in the range 1,027.1 – 1,027.7 kg/m^3.

4.1.2. Temperature and Salinity Flux Correlation Length Scales

As a separate question of spatial uncertainty, it is often useful to interpret the measured Lagrangian profiles in the context of surface fluxes from an external gridded data set. If a float is assigned an interpolated position, what is the uncertainty in these boundary values? Although any gridded surface data could be used, we have chosen the adjusted air-sea fluxes of heat, freshwater, and buoyancy from SOSE (section 2.4).

We estimate the correlation length scales for SOSE heat, buoyancy, and salinity fluxes by creating a grid of points within the domain at 3° intervals of latitude and longitude. Using the heat, buoyancy, and salinity flux time series at all of these grid points, we calculated several spatially lagged correlations with the time series record of all points within a higher-resolution grid of the Weddell Gyre separated by 0.5° intervals of latitude and longitude. The calculations had three time domains: the entire time series, austral summer (15 November to 15 March), and austral winter (15 May to 15 September). We assumed heat, salinity, and buoyancy fluxes have a similar decorrelation time scale as temperature and salinity (35 days); because of this, we were generous in our seasonal definitions (4 months for summer and winter) to include as many degrees of freedom as possible.

After the correlation calculations, we binned the results by distance in 5-km segments and calculated the average salinity, heat, and buoyancy flux correlation in each bin. Results are given in section 4.2.2.

4.1.3. Uncertainty Field and Representation Error Calculation

Finally, we used the calculated correlations to estimate the uncertainty (error) of a field. Our analysis is motivated by the intuition that if two locations are perfectly correlated, knowing a value at one location should allow you to know the value at the other. Similarly, if two locations are perfectly uncorrelated, knowing the value at one location yields no information about the value at the other location.

Expanding on this, let $T(x) \in \mathbb{R}$ be a random temperature or salinity field, as a function of location. Let $X \in \mathbb{R}^2$ be a random assumed location of the float, and let $X_{\text{true}} \in \mathbb{R}^2$ be the true location. Assuming that the field $T(x)$ changes slowly in x, we make a linear approximation at X_{true}, by defining the linear predictor:

$$L[T(X)|T(X_{\text{true}})] = E[T(X)] + \frac{cov[T(X), T(X_{\text{true}})]}{var(T(X_{\text{true}}))}[T(X_{\text{true}}) - E(T(X_{\text{true}})],$$

where $L[T(X)|T(X_{\text{true}})]$ is the linear predictor, $cov[T(X), T(X_{\text{true}})]$ is the covariance of the value of the field at the random location and the value of the field at the true location, $E(X)$ and $E(X_{true})$ are the expected values of the field at X and X_{true}, respectively, and $var(T)$ is the variance of the field. We can write the mean square error of this linear predictor:

$$\begin{aligned} \sigma^2_{\text{rep}} &= E[(T(X) - L)^2] = var(T) - \frac{cov[T(X), T(X_{\text{true}})]^2}{var(T)} \\ &= var(T)\left[1 - \frac{cov[T(X), T(X_{\text{true}})]^2}{var(T)^2}\right] \\ &= var(T)[1 - cor[T(X), T(X_{\text{true}})]^2]. \end{aligned} \tag{9}$$

σ_{rep} is the representation error. $cor[T(X), T(X_{\text{true}})]$ is the correlation of T between positions X and X_{true}. X can be distance in any metric—herein, we consider both latitude-longitude and PV coordinates. Notice that if $cor[T(X), T(X_{\text{true}})] = 1$, meaning the interpolated position is perfectly correlated with the true position, the measured value projects exactly onto the interpolated data and the total error will be only the measurement error. Conversely, if $cor[T(X), T(X_{\text{true}})] = 0$, the interpolated position is exactly uncorrelated with the interpolated position, and the measured value provides us no information about the value at the interpolated location. The error is then maximum and equal to the entire variance of the data set. This is the quantified uncertainty that matches our original intuition.

The total error, due to both representation error σ_{rep} and sensor error σ_{sensor}, arising from accuracy of the instrumental observation, is

$$\sigma_{\text{total}} = \sqrt{\sigma^2_{\text{sensor}} + \sigma^2_{\text{rep}}}. \tag{10}$$

where σ_{sensor} is reported from the manufacturer as 0.002 psu for salinity and 0.002 °C for temperature. Total error results are presented in the following sections.

4.2. Results

4.2.1. Temperature and Salinity Correlation Scales and Total Error

Temperature and salinity correlations are illustrated in Figures 9–12, along with the total error, which is the sum of representation error and sensor error (equation (10)). (The error field is discussed further in section 4.2.3, where it is shown that representation error dominates the total error, with sensor error much smaller.) Correlations in temperature and salinity as a function of depth and distance were calculated in 65 depth layers from the surface to 2,000 m and distances from 35 to 1,000 km in 35-km increments (Figures 9 and 10). Correlations and total error as a function of potential density and distance were calculated in four layers at 0.1 kg/m^3 increments from 27.3 to 27.7 kg/m^3 with the same distance binning as for depth (Figures 11 and 12).

The plotted correlations and errors are divided into two regimes: upper (0–300 m) and deeper (300–2,000 m) ocean. The upper layer encompasses the fresh, very cold surface layer affected directly by sea ice, down through the halocline to the temperature maximum/salinity maximum core of the upwelled North Atlantic Deep Waters. The deeper layer includes the rest of these northern waters and the top of the Weddell Sea Deep Water (Orsi et al., 1993).

All correlation calculations (Figures 9–12) show, unsurprisingly, that temperature/salinity correlation decays with increasing distance. Correlation length scales were longer in isopycnal layers (Figures 11 and 12) than in depth layers (Figures 9 and 10). This is likely a result of transient eddies and waves causing isopycnal heave or the sloping density structure of the gyre in the Weddell Gyre, which introduces noise and reduces correlation along depth surfaces. Correlation length scales were marginally longer in PV than Euclidean coordinates and with largely the same structure.

Total error within each layer increases with increasing distance, due to its dependence on correlation (equation (9)). Correlation scales were the longest in the upper ocean with significant correlation extending to 550 and 350 km in temperature and salinity, respectively (Figures 9a and 10a). The higher upper ocean

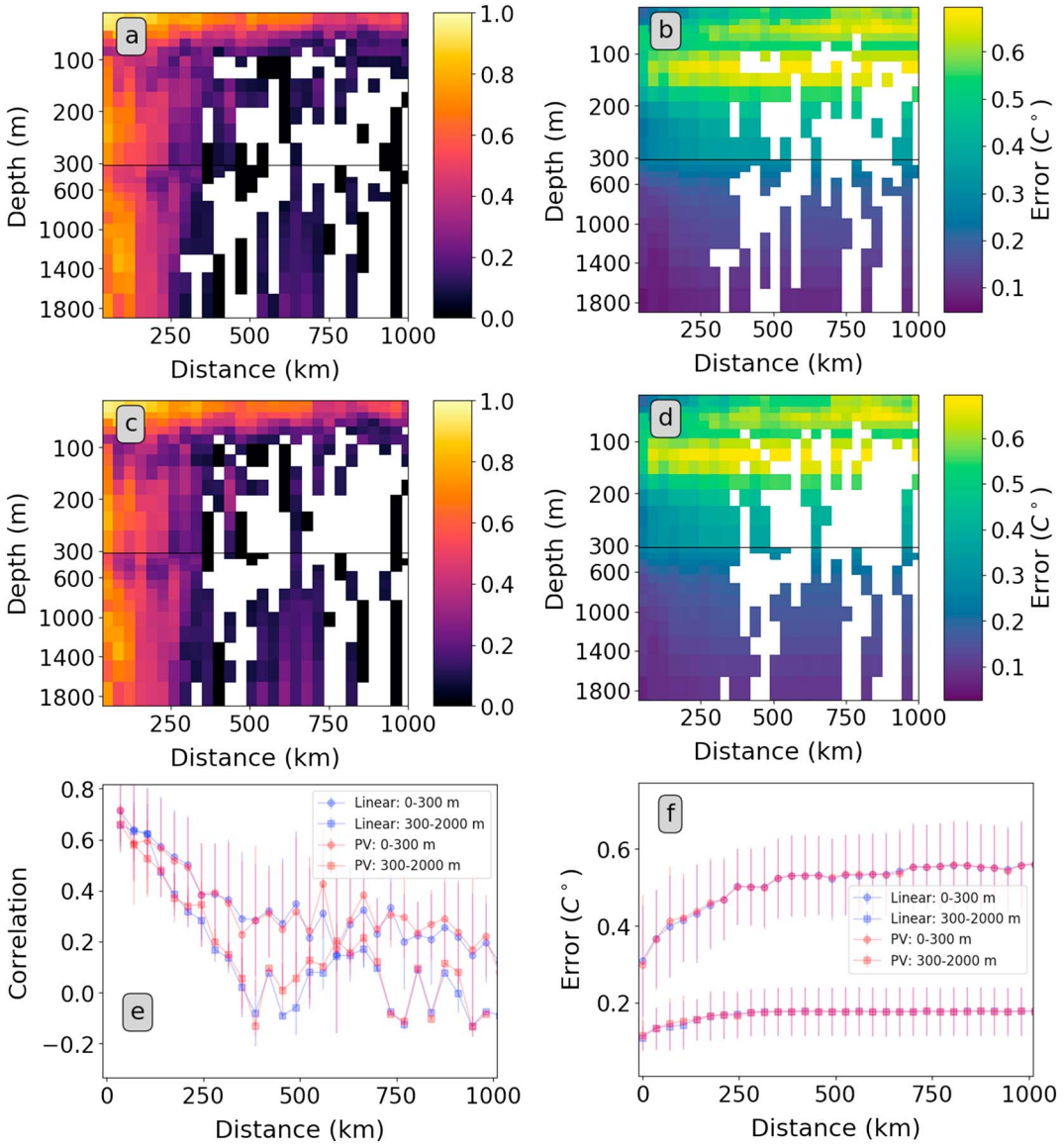

Figure 9. (a) Temperature correlation as a function of distance and depth at 90% confidence. (b) Total temperature error as a function of distance and depth at 90% confidence in regimes above and below 300 m. (c) Temperature correlation as a function of distance in PV coordinates and depth at 90% confidence. (d) Total temperature error as a function of distance in PV coordinates and depth at 90% confidence in regimes above and below 300 m. (e) Temperature correlation as a function of distance averaged over 0–300 m and over 300–2,000-m depth at 90% confidence; error bars represent standard deviation in correlation over these depth ranges. (f) Total temperature error as a function of distance at 90% confidence averaged over 0–300- and 300–2,000-m depth; error bars represent standard deviation in error over regime depths. PV = potential vorticity.

correlation compared with the deeper ocean is interpreted as resulting from the large spatial scale of air-sea buoyancy flux due to the annual cycle and longer time scale atmospheric variability.

Low correlations between 100 and 200 m are apparent in the full depth-resolution figures (Figures 9a, 9c, 10a, and 10c). This is the depth range of the surface mixed layer and halocline (Figure 4), and as explained above, correlations are low because of transient, localized heave of the highly stratified halocline.

Notably, the larger correlation scales in the upper ocean correspond to larger total error (Figures 9e, 9f, 10e, and 10f): the upper layer error is higher because the surface layer variance is high, while the deeper layer has more uniform temperature and salinity. We hypothesize that this is due to the wide range of overall variability in surface temperature and salinity and their fluxes during ice-free periods. The float time series example in

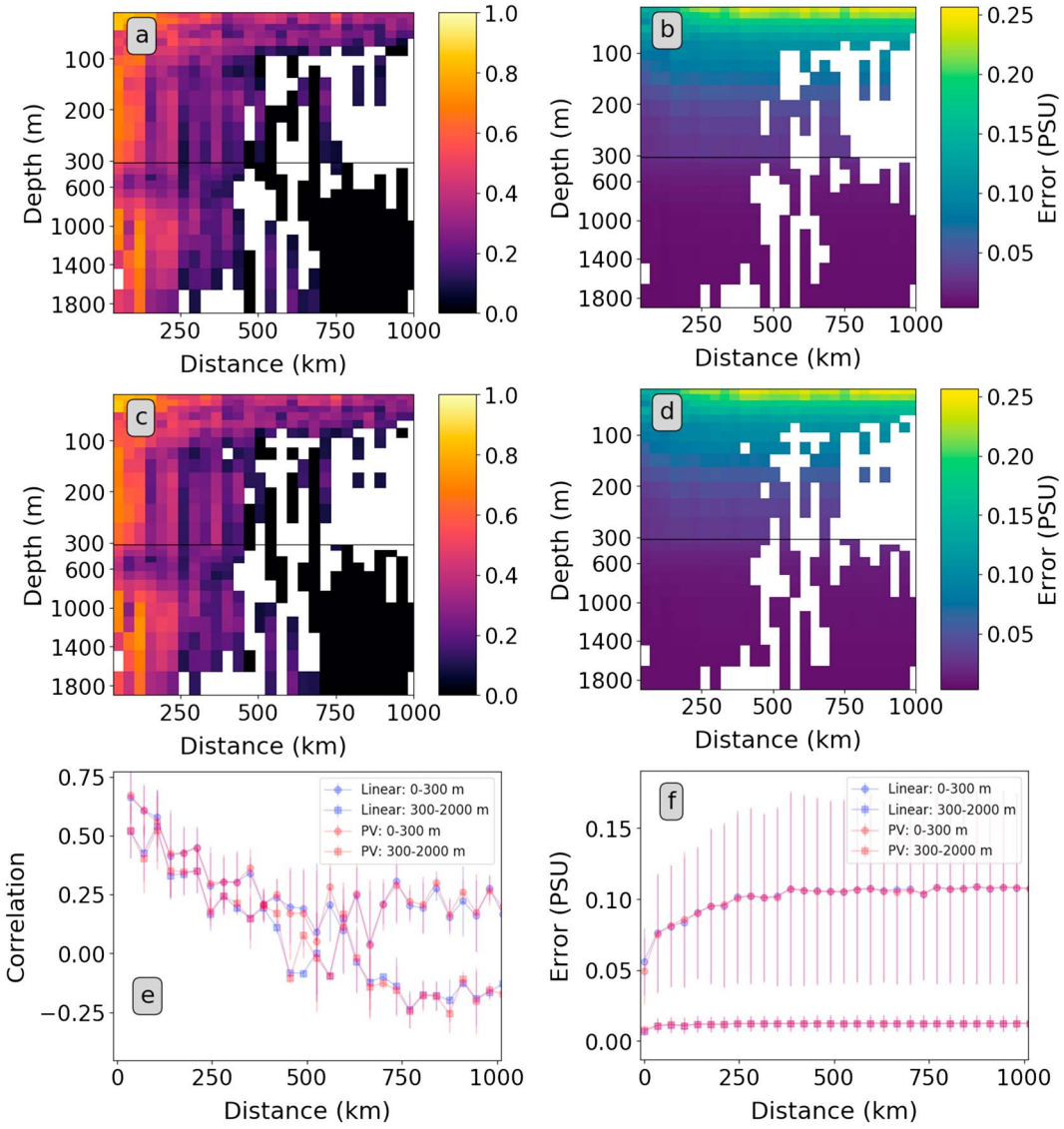

Figure 10. (a) Salinity correlation as a function of distance and depth at 90% confidence. (b) Total salinity error as a function of distance and depth at 90% confidence in regimes above and below 300 m. (c) Salinity correlation as a function of distance and depth at 90% confidence. (d) Total salinity error as a function of distance and depth at 90% confidence in regimes above and below 300 m. (e) Salinity correlation as a function of distance averaged over 0–300- and 300–2,000-m depth at 90% confidence; error bars represent standard deviation in correlation over these depth ranges. (f) Total salinity error as a function of distance at 90% confidence averaged over 0–300- and 300–2,000-m depth; error bars represent standard deviation in error over regime depths. PV = potential vorticity.

Figure 4 shows the strong seasonal dependence of the temperature, salinity, and mixed layer depth in the upper 200 m. This is a response to local surface forcing. An example of the along-track fluxes (from SOSE output) experienced by WMO ID # Argo Float 5901717 (Figure 13) shows large variations in heat and salinity flux during the ice-free seasons, with large spikes in fluxes during both ice formation and ice melt. The deeper layer does not have such large external forcing. Temperature and salinity variance in the lower layer is smaller, and hence, the total error is smaller.

4.2.2. Air-Sea Fluxes and Spatial Correlation Scales

SOSE surface fluxes into the ocean, along interpolated Argo tracks and RAFOS-enabled float tracks, are illustrated for float 5901717 (Figure 13). (Differences between the satellite-tracked and RAFOS-enabled float track fluxes during ice-free periods are solely due to the difference in sampling interval between 10-day satellite fixes and the daily acoustic positions; the fluxes match on the days of satellite fixes.) Fluxes were diagnosed from SOSE output.

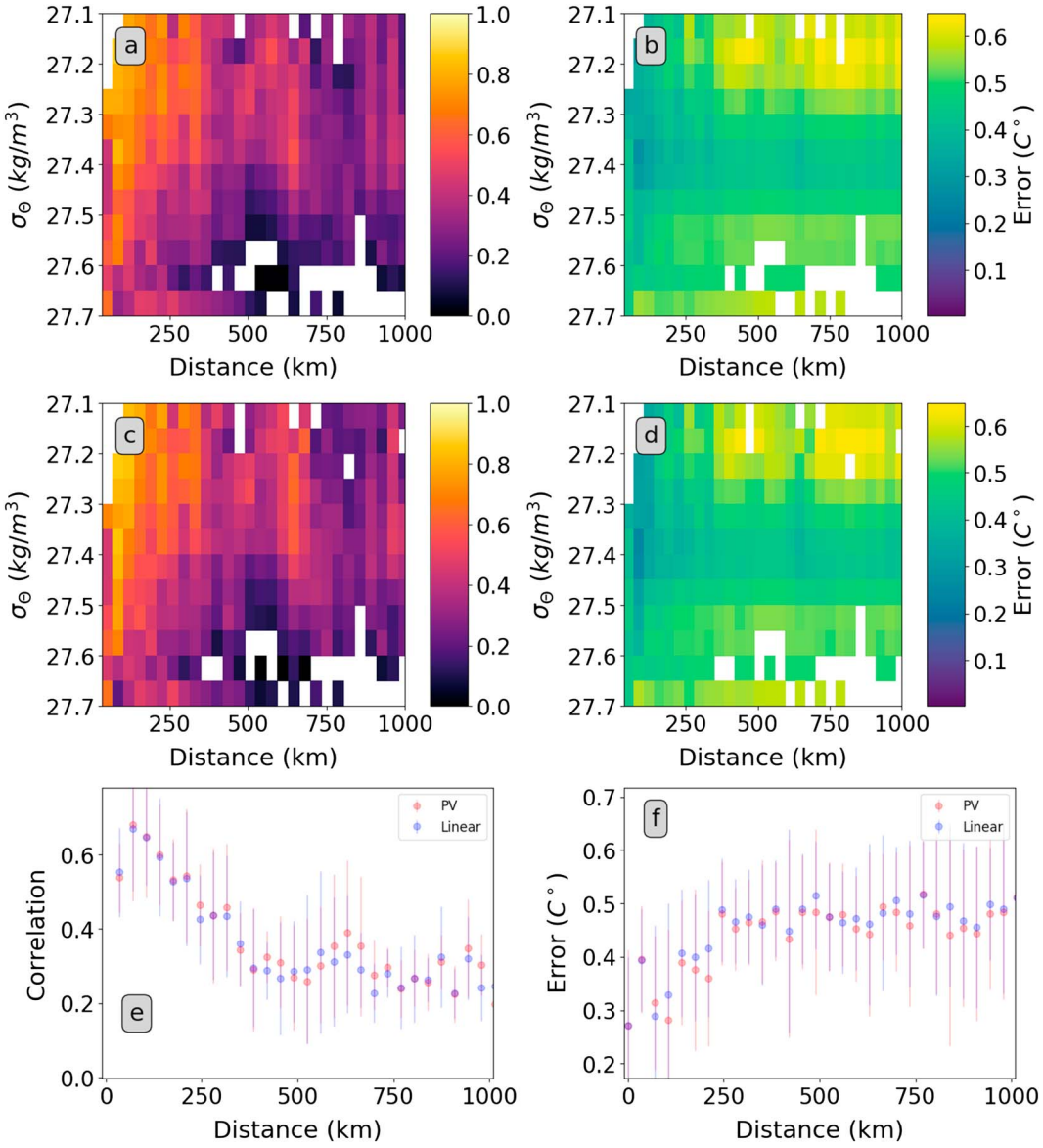

Figure 11. (a) Temperature correlation as a function of density and distance at 90% confidence. (b) Total temperature error as a function of distance and density at 90% confidence. (c) Temperature correlation as a function of density and distance in PV coordinates at 90% confidence. (d) Total temperature error as a function of distance in PV coordinates and density at 90% confidence. (e) Temperature correlation as a function of distance averaged over all density surfaces at 90% confidence; error bars represent standard deviation in correlation over all densities. (f) Total temperature error as a function of distance at 90% confidence averaged over all density surfaces; error bars represent standard deviation in error over all densities. PV = potential vorticity.

SOSE's under-ice heat flux into the ocean for float 5901717 is relatively constant with a mean of −6.77 W/m^2 (where positive flux heats the ocean). This trend of stable under ice heat fluxes is seen over all floats. The average under-ice heat flux over the entire data set was −8.06 W/m^2. At the end of periods of position loss, we observe a flux of fresh water into the ocean, closely followed by positive heat flux into the ocean which peaks close to 200 W/m^2. The time lag between the significant spikes in heat and salinity flux can be explained by the seasonal cycle of sea ice formation and melt. At the end of the sea ice season, the energy that would be entering the ocean due to increased solar forcing is instead stored in latent heat and used to melt sea ice. Once a significant fraction of the ice has melted, then heat enters the ocean as seen by the large spike in positive heat flux and the observed time lag.

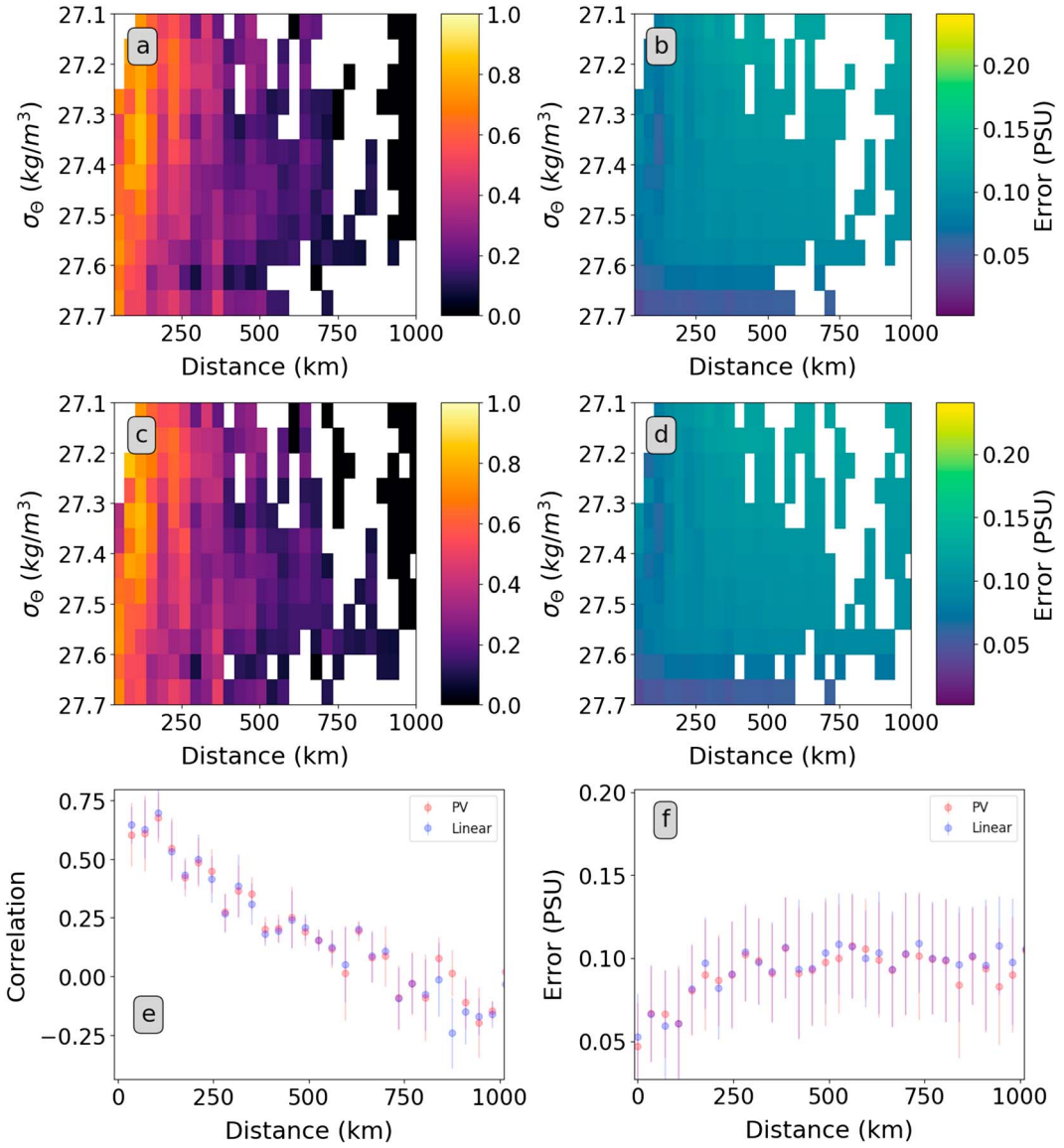

Figure 12. (a) Salinity correlation as a function of density and distance at 90% confidence. (b) Total salinity error as a function of distance and density at 90% confidence. (c) Salinity correlation as a function of density and distance in PV coordinates at 90% confidence. (d) Total salinity error as a function of distance in PV coordinates and density at 90% confidence. (e) Salinity correlation as a function of distance averaged over all density surfaces at 90% confidence; error bars represent standard deviation in correlation over all densities. (f) Total salinity error as a function of distance at 90% confidence averaged over all density surfaces; error bars represent standard deviation in error over all densities. PV = potential vorticity.

Conversely, at the beginning of the ice season, there is a significant flux of heat out of the ocean (on the order of −200 W/m^2) closely followed by a salinity flux into the ocean. This is an example of seasonal forcing causing sea ice formation followed by brine rejection. The sea ice, once formed, acts as a blanket on the ocean and prevents significant air-sea flux.

We have plotted the correlations and related uncertainties of mean under-ice surface fluxes into and out of the ocean for heat, salinity, freshwater, and buoyancy as a function of distance (Figure 14). Similar to float salinity and temperature measurements, surface fluxes at close separations are more correlated than locations that are far apart. We observe that overall buoyancy flux is dominated by the salinity flux signal for these Weddell Gyre floats. For 8 months of position loss, we calculate maximum heat flux, buoyancy flux, and salinity flux errors to be 28.5 W/m^2, 212.4 W/m^2, and 1.8×10^{-3} kg · m^{-2} · s^{-1}, respectively. Heat flux correlations are independent of season. In contrast, salinity correlations are about 0.2 higher in summer than in winter; buoy-

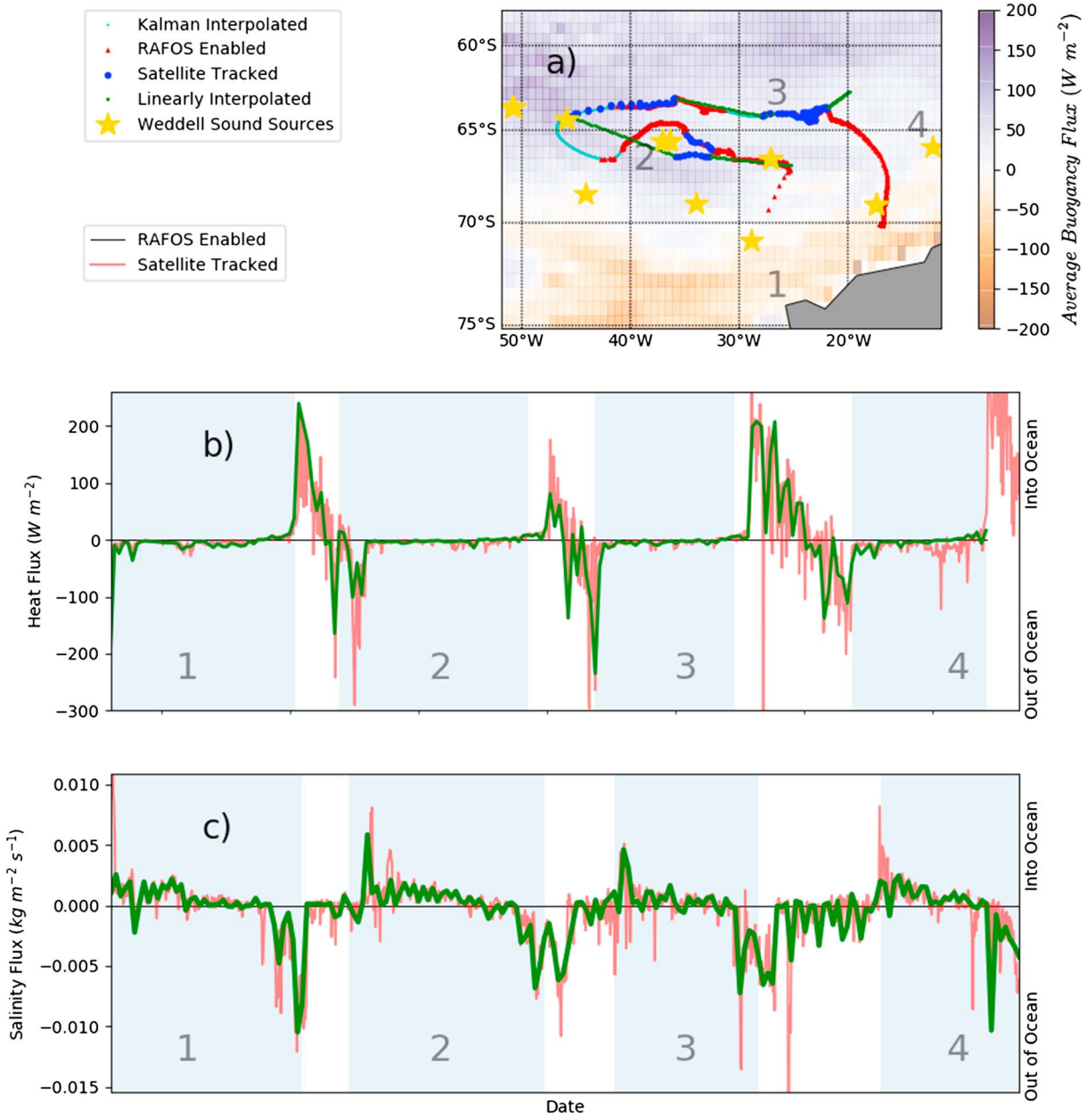

Figure 13. Comparison of RAFOS-enabled float and interpolated Argo float for float WMO5901717. (a) Map of tracks. Acoustically tracked float locations (red dots), satellite-tracked float locations (blue dots), linearly interpolated location (green dots), and acoustic sound sources (stars). Shading is buoyancy flux, temporally averaged over the entire SOSE record, reported in units of equivalent heat flux (W/m^2). (b) Surface heat flux along the RAFOS-enabled (red) and interpolated (green) Argo float tracks. (c) Surface salinity flux along the RAFOS-enabled (red) and interpolated (green) Argo float tracks. Gray shading indicates periods of position loss due to sea ice, also labeled on the map.

ancy correlations follow salinity. This seasonal change in correlation results in a seasonal variation in the flux uncertainty of buoyancy and heat of 34.2% and 27.2%, respectively. As seen in Figure 13, flux regimes depend strongly on season and ice conditions. The difference between summer and winter salinity correlations is due to greater modeled regional structure during ice formation (and hence brine rejection) than to melting.

The SOSE flux results can also be used to examine our assumption of temporal invariability of correlation scales of upper ocean properties—we hypothesize that correlation length scales of the upper ocean heat and salinity should correspond with associated heat and salinity flux correlations. In Figure 14, we compare the qualitative and quantitative structures of these calculated values by overlaying the ocean temperature and salinity correlation averaged over the upper 50 m (described in section 4.2) on the heat and salinity flux

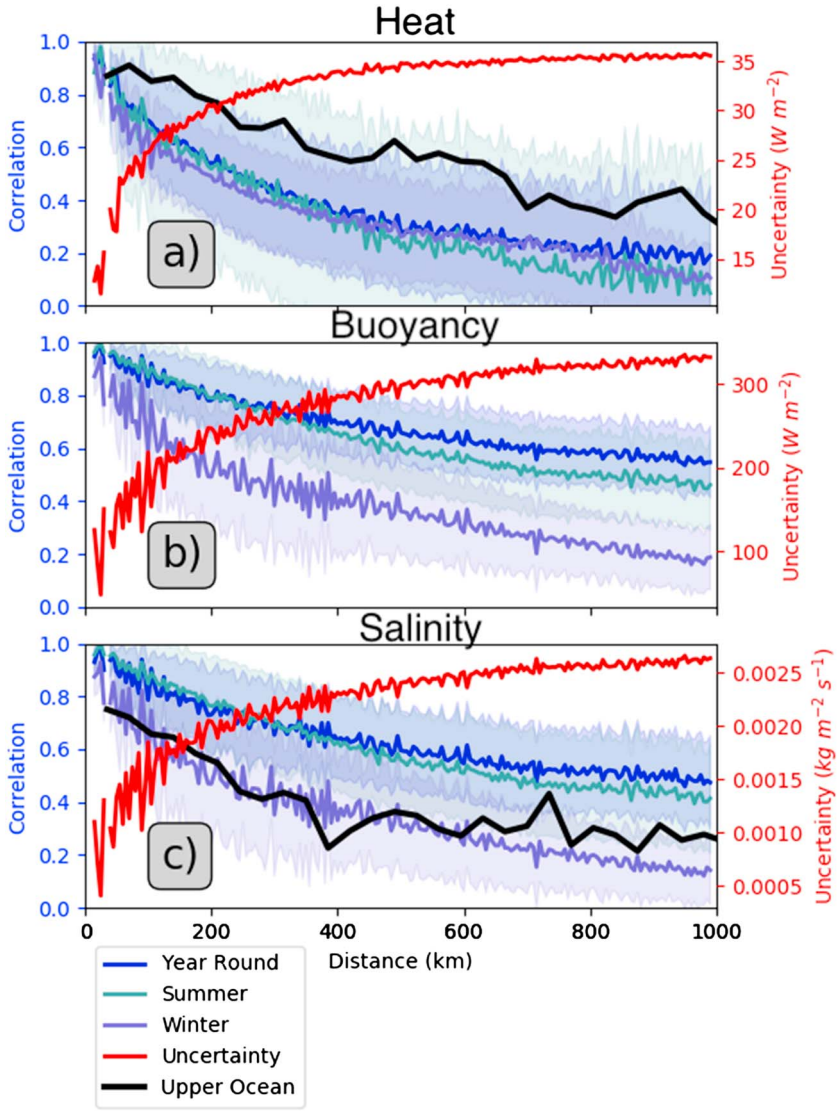

Figure 14. Spatial correlation of surface fluxes from Southern Ocean State Estimate for (a) heat (b) buoyancy, and (c) salinity, for winter (15 May to 15 September), summer (15 November to 15 March), and year round. Colored swath around lines represent standard deviation of data in each bin. Red line is uncertainty calculated from year-round correlation curve. Black line is upper ocean temperature and salinity correlation averaged over top 50 m (from Figures 9 and 10).

plots, respectively. We observe that temperature correlation is about 0.1 greater than heat flux correlation and salinity correlation matches well with wintertime salinity flux correlation. One component of upper ocean temperature or salinity is the time integral of surface heat or salinity flux, so we hypothesize that the upper ocean smooths high-frequency flux components of the heat and salinity flux signal. This could result in longer correlation scales for upper ocean temperature and salinity than for heat and salinity flux. It is problematic then that upper ocean salinity correlation closely matches wintertime salinity flux correlation because we would expect it to be higher and because much of the satellite-tracked float data were collected in the summertime. This suggests that the modeled correlation scales of heat and salinity flux in the Weddell Sea region of SOSE may be inaccurate and should be treated with some skepticism.

4.2.3. Salinity and Temperature Error as a Function of Position Loss Time

We have estimated position uncertainty in latitude-longitude as a function of time (section 3) and temperature and salinity uncertainty as a function of latitude-longitude distance (section 4). In this section, we combine these two calculations to estimate representation error as a function of position loss. These uncertainty esti-

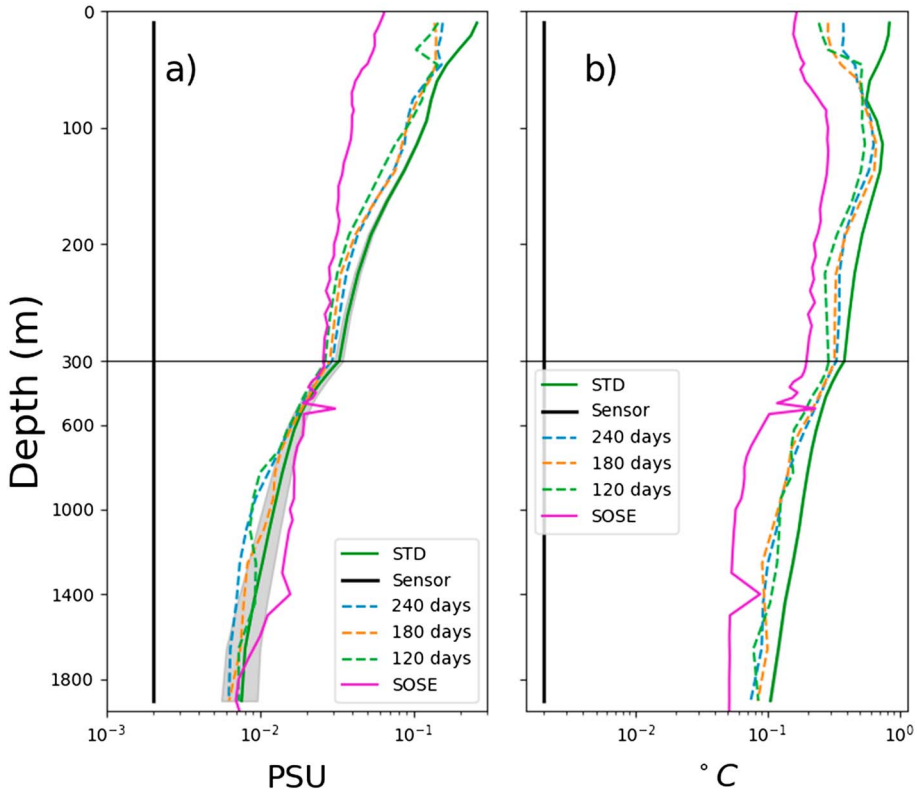

Figure 15. (a) Salinity and (b) temperature errors in the Weddell Sea as a function of depth and time of position loss. Solid green curve represents standard deviation (STD) of temperature and salinity of all satellite-tracked float profiles. Shaded region on either side of the green curve is the sensor error. Dashed lines are measurement uncertainty for various durations of position loss. Cyan curve is average of SOSE model uncertainty in the study area. Black lines represent sensor uncertainty: 0.002 psu and 0.002 °C. SOSE = Southern Ocean State Estimate.

mates are useful for synthesizing Lagrangian data of unknown position into gridded Eulerian products and are calculated for the first time.

Maximum total error in temperature and salinity, as a function of depth for a location to which an under-ice float profile has been assigned through interpolation, is estimated as follows: given a specified TIPT, maximum position uncertainty is estimated using Figure 7 (section 3). The temperature and salinity error profiles are then taken from the temperature and salinity error maps (Figures 9d and 10d). For example, the column between 70 and 105 km would be used for a position uncertainty of 90 km.

The maximum total temperature and salinity error due to linear interpolation of positions for 120, 180, and 240 days of position loss as well as the standard deviation of the Weddell Gyre Argo float temperature and salinity data are shown in Figure 15. For the upper ocean (shallower than 300 m), maximum error can equal 0.66 °C or 94.5% of the total variance in the temperature field and 0.15 psu or 91.4% of the variance in salinity. Gridded products that do not account for this linear interpolation error in position measurement will project this uncertainty onto model error. The prescribed model error for SOSE (Mazloff et al., 2010) averaged over our study area is also shown in Figure 15. In the Weddell Gyre, SOSE uses a maximum model error of 0.24 °C in temperature and 0.067 psu in salinity. Comparing prescribed model error with maximum linear interpolation uncertainties, we find that linear interpolation uncertainty is 276% of SOSE model error in temperature and 229% in salinity. These representation errors are not only significant in the upper ocean. For the deeper parts of Argo float profiles (300–2,000 m), we find that uncertainty due to linear interpolation can be as high as 0.16 °C, or 55.9% of the total variance in the temperature field, and 0.01 psu, or 91.1% of the signal in salinity.

In the broader sense, this highlights the importance of this type of accounting for use in all applications. If higher uncertainty due to interpolated positions is not combined with measurement uncertainty, any product generated is likely to overfit the data. The Southern Ocean contributes to the long-term oceanic heat gain. Desbruyères et al. (2016) found a warming trend in the deep ocean below 2,000 m of $2.17 \pm 0.70 \times 10^{-3}$ °C/year

averaged from 1991 to 2010. At 2,000 m, we estimate a temperature uncertainty of 8.49×10^{-2} °C, which represents 39.1 years of warming at the Desbruyères et al., 2016 rate. This comparison might be misleading because of the different depth ranges, given that the warming trend has been minimal in the Southern Ocean's Deep Water layer between the thermocline and abyssal water (Purkey & Johnson, 2010). Nevertheless, the reader is cautioned that as autonomous technologies advance and the Argo array heads into the abyss with the deployment of Deep Argo, linear interpolation uncertainty, if inappropriately accounted for, could mask subtle residual signals.

A noteworthy point is that the insonification of the Weddell Gyre and use of RAFOS-enabled profiling floats eliminates linear interpolation error for the temperature and salinity profiles when RAFOS tracking is possible. In our study, RAFOS-enabled floats received satellite positioning on 4% of days and were acoustically tracked on 56% of days. As noted in Klatt et al., 2007, sound transmission in the Weddell Sea is more efficient during ice-free conditions, and consequently, floats were acoustically tracked on 52% of days while under the ice and 61% of days while out of ice. A denser network of sound sources or improvements in under-ice sound transmission or detection could improve this tracking rate. Any long-term study of under ice processes should consider acoustic tracking to reduce representation error due to linear interpolation.

5. Conclusions

Position uncertainty and spatial correlation scales for temperature and salinity were combined to produce error estimates as a function of time-of-position-loss for a set of Argo floats deployed in the Weddell Gyre. Position uncertainty was calculated by linearly interpolating float track data, in latitude-longitude and PV coordinates, at intervals from 1 to 8 months. Float positions were derived from satellite-tracked Argo floats, RAFOS-enabled Argo floats (when available), and a SOSE particle release simulation. The position uncertainty estimated from the RAFOS-enabled float tracks for 8 months of ice cover was 116 ± 148 km. The estimates derived from satellite-tracked and RAFOS-enabled data sets were in good agreement, although there were only enough satellite-tracked data during ice-free periods to make estimates for 4 months of position loss. PV interpolation reduced position uncertainty for all data sets. We interpreted the quadratic structure of the position loss as a function of time using a Brownian Bridge, which assumes that the deviation from the linearly interpolated track is due to random motion resulting from the mesoscale eddy field. Maximum position uncertainty was similar during summer and winter which suggests that the parametrized diffusivity (σ^2) is temporally invariant.

The agreement between position uncertainties estimated from satellite-tracked floats (during ice-free periods of continuous 10-day sampling) and RAFOS-tracked floats, as well as the similarities of winter and summer position uncertainty statistics, suggests that satellite-tracked profiling floats can be used to estimate position uncertainties in other parts of the seasonally ice-covered Southern Ocean that have no RAFOS tracking. In contrast, particle tracking in the 0.16° SOSE iteration yielded position uncertainties that were approximately 75% of the observation-based method, possibly because SOSE at this spatial resolution underestimates eddy energy. Therefore, we cannot recommend using the 0.16° SOSE to estimate these statistics throughout the Southern Ocean.

Within the range of 1- to 8-month position uncertainty, maximum error in salinity and temperature above 300 m was 0.15 psu and 0.66 °C, respectively; maximum uncertainty in temperature and salinity between 300 and 2,000 m was 0.01 psu and 0.16 °C, respectively. Spatial correlation was longer along isopycnal surfaces and in the mixed layer. Error due to linear interpolation of float positions increased as a function of time of position loss and was up to 81.9% of the variance in Weddell Gyre float temperature measurements and 96.5% of the variance in Weddell Gyre float salinity measurements for 8 months of position loss. Temperature and salinity uncertainties are important inputs to state estimates, which are constrained by observations. We recommend that temperature and salinity uncertainties in state estimates be increased above measurement error for Argo floats that are under the ice.

We also found maximum heat flux, salinity flux, and buoyancy flux errors to be 28.5 W/m^2, 1.8×10^{-3}kg · m^{-2} · s^{-1}, and 212.4 W/m^2, respectively. Heat flux errors had little seasonal variability. Salinity, freshwater, and buoyancy flux correlation values were approximately 0.2 greater during austral summer than in austral winter. Spatial correlation scales in temperature and salinity were calculated along both depth and density surfaces and presented as a function of total distance.

If an observing system aims to resolve spatial variations in upper ocean temperature and salinity, acoustic tracking is recommended. The RAFOS float tracking used for the Weddell Gyre floats here was carried out with a new approach that we developed in the course of this study, using a Kalman smoother, to be described in a subsequent publication. This approach reduces the tracking error. These errors are not treated here; rather, the tracking is treated as if it is exact.

As the volume of biogeochemical data increases, our uncertainty analysis can be repeated for variables other than temperature and salinity. For a biogeochemical tracer, such as alkalinity, which is likely to have distributions that are similar to salinity (Lee et al., 2006), we would expect correlation scales to be similar to those for salinity. Biogeochemical tracers that are closely linked to surface ventilation and seasonal blooms will likely have correlation scales and uncertainties that differ from temperature and salinity and will need to be calculated from in situ data sets, which are as yet inadequate.

Acknowledgments
Acoustic TOA processing for the RAFOS floats was carried out by Dr Cathrine Hancock. Dr Bruce Cornuelle suggested and assisted with development of the Kalman smoother algorithm for RAFOS-tracking. Drs Sarah Gille and Ivana Cerovecki provided insightful comments. The authors also thank two anonymous reviewers for their thoughtful and detailed edits. This work was supported by the SOCCOM project under NSF Award PLR-1425989. K. Speer acknowledges support from NSF OCE 1231803. S. Riser acknowledges support from NOAA grants NA17RJ1232 and NA15OAR4320063 and NSF grant OPP1429342. Argo data sources are listed in the references. Our quality control of the float locations revealed issues that have been reported to Argo and subsequently corrected. The RAFOS float trajectory data are preliminary and available on request from the authors. SOSE output is publicly available for download at http://sose.ucsd.edu/.

References

Abernathey, R., Cerovecki, I., Holland, P., Newsom, E., Mazloff, M., & Talley, L. (2016). Southern Ocean water mass transformation driven by sea ice. *Nature Geoscience*. https://doi.org/10.1038/ngeo2749

Antonov, J., Seidov, D., Boyer, T., Locarnini, R., Mishonov, A., Garcia, H., et al. (2010). World Ocean Atlas 2009. In S. Levitus (Ed.), *Salinity* (Vol. 2, pp. 184). Washington, DC: US Government Printing Office.

Balwada, D., Speer, K., LaCasce, J., Owens, B., Marshall, J., & Raffaele, F. (2016). Circulation and stirring in the southeast Pacific Ocean and the Scotia Sea sectors of the Antarctic Circumpolar Current. *Journal of Physical Oceanography*, *46*(7), 2005–2027.

Boebel, O. (2009). The expedition of the research vessel" polarstern" to the Antarctic in 2008/2009 (ANT-XXV/2). Berichte zur Polar-und Meeresforschung (Reports on Polar and Marine Research) 595.

Carval, T., Keeley, R., Takatsuki, Y., Yoshida, T., Schmid, C., Goldsmith, R., et al. (2015). Argo user's manual v3. 2.

Cerovečki, I., & Mazloff, M. R. (2016). The spatiotemporal structure of diabatic processes governing the evolution of Subantarctic Mode Water in the Southern Ocean. *Journal of Physical Oceanography*, *46*(2), 683–710.

Cerovečki, I., Talley, L. D., Mazloff, M. R., & Maze, G. (2013). Subantarctic Mode Water formation, destruction, and export in the eddy-permitting Southern Ocean State Estimate. *Journal of Physical Oceanography*, *43*(7), 1485–1511.

Cerovečki, I., Talley, L. D., & Mazloff, M. R. (2011). A comparison of Southern Ocean air–sea buoyancy flux from an ocean state estimate with five other products. *Journal of Climate*, *24*(24), 6283–6306.

D'Ortenzio, F., Lavigne, H., Besson, F., Claustre, H., Coppola, L., Garcia, N., et al. (2014). Observing mixed layer depth, nitrate and chlorophyll concentrations in the northwestern Mediterranean: A combined satellite and NO_3 profiling floats experiment. *Geophysical Research Letters*, *41*, 6443–6451. https://doi.org/10.1002/2014GL061020

Desbruyères, D. G., Purkey, S. G., McDonagh, E. L., Johnson, G. C., & King, B. A. (2016). Deep and abyssal ocean warming from 35 years of repeat hydrography. *Geophysical Research Letters*, *43*, 10,356–10,365. https://doi.org/10.1002/2016GL070413

Fahrbach, E., & de Baar, H. (2010). The expedition of the research vessel" polarstern" to the Antarctic in 2008 (ANT-XXIV/3). Berichte zur Polar-und Meeresforschung (Reports on Polar and Marine Research) 606.

Fahrbach, E., Hoppema, M., Rohardt, G., Boebel, O., Klatt, O., & Wisotzki, A. (2011). Warming of deep and abyssal water masses along the Greenwich meridian on decadal time scales: The Weddell Gyre as a heat buffer. *Deep Sea Research Part II: Topical Studies in Oceanography*, *58*(25), 2509–2523.

Fetterer, F., Knowles, K., Meier, W., & Savoie, M. (2016). Sea Ice Index Version 2s. https://doi.org/10.7265/N5736NV7

Gille, S. T., & Kelly, K. A. (1996). Scales of spatial and temporal variability in the Southern Ocean. *Journal of Geophysical Research*, *101*, 8759–8773.

Gille, S. T., & Romero, L. (2003). Statistical behavior of ALACE floats at the surface of the Southern Ocean. *Journal of Atmospheric and Oceanic Technology*, *20*(11), 1633–1640.

Gray, A. R., & Riser, S. C. (2014). A global analysis of Sverdrup balance using absolute geostrophic velocities from Argo. *Journal of Physical Oceanography*, *44*(4), 1213–1229.

Haumann, F. A., Gruber, N., Münnich, M., Frenger, I., & Kern, S. (2016). Sea-ice transport driving Southern Ocean salinity and its recent trends. *Nature*, *537*(7618), 89–92. https://doi.org/10.1038/nature19101

Johnson, K. S., & Claustre, H. (2016). Bringing biogeochemistry into the Argo age. *Eos*, *97*. https://doi.org/10.1029/2016EO062427

Jones, E. M., Bakker, D. C., Venables, H. J., & Hardman-Mountford, N. J. (2015). Seasonal cycle of CO_2 from the sea ice edge to island blooms in the Scotia Sea, Southern Ocean. *Marine Chemistry*, *177*, 490–500.

Klatt, O., Boebel, O., & Fahrbach, E. (2007). A profiling float's sense of ice. *Journal of Atmospheric and Oceanic Technology*, *24*(7), 1301–1308.

LaCasce, J., Ferrari, R., Marshall, J., Tulloch, R., Balwada, D., & Speer, K. (2014). Float-derived isopycnal diffusivities in the DIMES experiment. *Journal of Physical Oceanography*, *44*(2), 764–780. https://doi.org/10.1175/JPO-D-13-0175.1

LaCasce, J., & Speer, K. (1999). Lagrangian statistics in unforced barotropic flows. *Journal of Marine Research*, *57*(2), 245–274.

Lee, K., Tong, L. T., Millero, F. J., Sabine, C. L., Dickson, A. G., Goyet, C., et al. (2006). Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans. *Geophysical Research Letters*, *33*, L19605. https://doi.org/10.1029/2006GL027207

Locarnini, R., Mishonov, A., Antonov, J., Boyer, T., Garcia, H., Baranova, O., et al. (2010). World Ocean Atlas 2009. Vol. 1, Temperature, NOAA Atlas NESDIS, 68.

Locarnini, R., Mishonov, A., Antonov, J., Boyer, J., Garcia, H., Baranova, O., et al. (2013). Temperature NOAA Atlas NESDIS 73. *World Ocean Atlas*, *1*, 40.

Mazloff, M. R., Heimbach, P., & Wunsch, C. (2010). An eddy-permitting Southern Ocean state estimate. *Journal of Physical Oceanography*, *40*(5), 880–899.

McDougall, T. J., & Barker, P. M. (2011). Getting started with TEOS-10 and the Gibbs Seawater (GSW) oceanographic toolbox. *SCOR/IAPSO WG*, *127*, 1–28.

Meredith, M. P., Meijers, A. S., Naveira Garabato, A. C., Brown, P. J., Venables, H. J., Abrahamsen, E. P., et al. (2015). Circulation, retention, and mixing of waters within the Weddell-Scotia Confluence, Southern Ocean: The role of stratified Taylor columns. *Journal of Geophysical Research: Oceans*, *120*, 547–562. https://doi.org/10.1002/2014JC010462

Nguyen, A., Ocaña, V., Garg, V., Heimbach, P., Toole, T., Krishfield, R., et al. (2017). On the benefit of current and future ALPS data for improving Arctic coupled ocean-sea ice state estimation. *Oceanography*, *30*(2), 69–73. https://doi.org/10.5194/os-12-1-2016

Ninove, F., Traon, P.-Y. L., Remy, E., & Guinehut, S. (2016). Spatial scales of temperature and salinity variability estimated from Argo observations. *Ocean Science*, *12*(1), 1–7. https://doi.org/10.5194/os-12-1-2016

Orsi, A. H., Nowlin, W. D., & Whitworth, T. (1993). On the circulation and stratification of the Weddell Gyre. *Deep Sea Research Part I: Oceanographic Research Papers*, *40*(1), 169–203.

Purkey, S. G., & Johnson, G. C. (2010). Warming of global abyssal and deep Southern Ocean waters between the 1990s and 2000s: Contributions to global heat and sea level rise budgets. *Journal of Climate*, *23*(23), 6336–6351.

Purkey, S. G., & Johnson, G. C. (2012). Global contraction of Antarctic Bottom Water between the 1980s and 2000s. *Journal of Climate*, *25*(17), 5830–5844.

Rauch, H. E., Striebel, C., & Tung, F. (1965). Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, *3*(8), 1445–1450.

Reeve, K., Boebel, O., Kanzow, T., Strass, V., Rohardt, G., & Fahrbach, E. (2015). Objective mapping of Argo data in the Weddell Gyre: A gridded dataset of upper ocean water properties. *Earth System Science Data Discussions*, *8*(1).

Revuz, D., & Yor, M. (2013). *Continuous Martingales and Brownian motion* (Vol. 293). Springer Science & Business Media. https://doi.org/10.1007/978-3-662-06400-9

Rhein, M., Rintoul, S., Aoki, S., Campos, E., Chambers, D., Feely, R., et al. (2013). Observations: Ocean. *Climate Change*, 255–316.

Rossby, T., Dorson, D., & Fontaine, J. (1986). The RAFOS system. *Journal of Atmospheric and Oceanic Technology*, *3*(4), 672–679.

SEANOE (2000). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). https://doi.org/10.17882/42182

Sarmiento, J., Gruber, N., Brzezinski, M., & Dunne, J. (2004). High-latitude controls of thermocline nutrients and low latitude biological productivity. *Nature*, *427*(6969), 56–60.

Wang, J., Mazloff, M. R., & Gille, S. T. (2014). Pathways of the Agulhas waters poleward of 29°. *Journal of Geophysical Research: Oceans*, *119*, 4234–4250. https://doi.org/10.1002/2014JC010049

Wong, A. P., & Riser, S. C. (2011). Profiling float observations of the upper ocean under sea ice off the Wilkes Land Coast of Antarctica. *Journal of Physical Oceanography*, *41*(6), 1102–1115.

Wooding, C. M., Furey, H. H., & Pachece, M. A. (2005). RAFOS float processing at the Woods Hole Oceanographic Institution. Tech. rep. DTIC Document.

Zweng, M., Reagan, J., Antonov, J., Locarnini, R., Mishonov, A., Boyer, T., et al. (2013). World Ocean Atlas 2013, volume 2: Salinity. *NOAA Atlas NESDIS*, *74*, 39.